



**ANA LAGES COVILHÃ ANÁLISE DE CONGLOMERADOS E ESTATÍSTICAS
CORTEZ DAS SOCIEDADES NÃO FINANCEIRAS**



**ANA LAGES COVILHÃ
CORTEZ**

**ANÁLISE DE CONGLOMERADOS E ESTATÍSTICAS
DAS SOCIEDADES NÃO FINANCEIRAS**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Matemática e Aplicações, realizada sob a orientação científica da Doutora Adelaide Freitas, Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro.

Dedico este trabalho a minha adorada mãe Amélia de que muito me orgulho, que não pode estar presente, para comigo partilhar este momento inesquecível, de modesta conquista.

Aos meus filhos Orfeu e Jaciana, razão da minha existência, que apesar da distância que nos separa souberam estar próximo de mim, com palavras de carinho. Foi difícil ficar longe de vós, mas hoje é reconfortante, olhar para trás e fazer uma retrospectiva do caminho percorrido.

Aos meus irmãos, que são o meu suporte e que em todos momentos, estiveram presentes com uma palavra de encorajamento e consolo.

o júri

presidente	Professor Doutor Pedro Filipe Pessoa Macedo, Professor Auxiliar, Universidade de Aveiro
vogal-arguente principal	Professora Doutora Anabela Virgínia dos Santos Flores da Rocha, Professora Adjunta, Universidade de Aveiro
vogal-orientador	Professora Doutora Adelaide de Fátima Baptista Valente Freitas, Professora Auxiliar, Universidade de Aveiro

agradecimentos

Agradecemos à administração do Banco Nacional de Angola, que tornou possível a realização deste Mestrado em Matemática e Aplicações, na especialização de Estatística e Otimização, bem como a todos os responsáveis e colegas dos Departamentos de Recursos Humanos, Controlo Cambial, Estatística, Gabinete do Governador e Relações Públicas do mesmo Banco que, de forma direta ou indireta, contribuíram para o culminar desta tarefa árdua.

Ao Banco de Portugal, que autorizou a realização do estágio curricular no Departamento de Estatística, na área da Central de Balanços, sobre o tema “Estatísticas de Sociedades Não Financeiras de Angola”. Em especial agradecemos à Direção do referido Departamento, coordenadores de núcleos e todos quantos se mostraram disponíveis para colaborar e transmitir os seus conhecimentos.

Agradecemos à orientadora, Doutora Adelaide Freitas, Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro, pela sua disponibilidade em acompanhar a minha evolução desde o primeiro ano do mestrado, bem como o apoio incondicional prestado nesta fase final e crucial. Reiteramos, ainda os nossos agradecimentos pela sua orientação, levada de forma moderada e firme, assim como a sua sabedoria na transmissão dos conhecimentos científicos, que serviram de suporte à preparação deste trabalho.

Reiteramos os agradecimentos à Doutora Isabel Pereira, Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro por todo o apoio prestado, incluindo no tratamento de aspetos administrativos inerentes ao ato de matrícula e aceitação na Universidade de Aveiro e pelo voto de confiança depositado em mim.

Reiteramos os nossos agradecimentos ao quadro de professores do Departamento de Matemática da Universidade de Aveiro, bem como à área administrativa, pela permanente disponibilidade e atenção durante os dois anos que durou o Mestrado em Matemática e Aplicações, na especialização em Estatística e Otimização.

Ainda aos colegas que sempre se mostraram disponíveis para partilhar os seus conhecimentos e experiências no mundo da investigação científica.

À minha querida e inesquecível mãe, cuja força batalhadora e perseverança, conseguiu transmitir aos seus filhos herdeiros dos princípios que sempre primaram a sua ideologia de vida.

À minha família, em especial aos meus irmãos e aos meus filhos que, mesmo à distância, souberam deixar palavras de incentivo, carinho e apreço, em todos os momentos deste percurso.

A Deus que guiou e orientou os meus passos, aclarou as minhas ideias e iluminou o meu caminho, tornando-o abençoado.

palavras-chave

Angola; análise de conglomerados; Banco de Portugal; método Cold-Deck; quadro de setor; setor não petrolífero; Sociedades Não Financeiras.

resumo

Na primeira parte do estudo realiza-se a análise das medidas de similaridade e dissimilaridades entre os 12 setores de atividade do setor não petrolífero (excetuando o setor financeiro), que permitiu identificar conglomerados, ou clusters, de setores com comportamento e evolução semelhantes entre si e dissemelhantes face aos outros grupos.

A análise de conglomerados permitiu observar a existência de ligações e características comuns entre os diferentes setores de atividade, revelando dinâmicas setoriais que podem ser potenciadas no sentido de introduzir a competitividade e a diversificação da economia angolana.

Sendo a construção um setor organizado e com boas perspetivas de crescimento, este foi selecionado para objeto de estudo de caso prático, tendo, para tal, sido aprofundado o conhecimento metodológico e funcional da formulação das estatísticas das Sociedades Não Financeiras.

Assim, o caso prático pretendeu a aplicação da metodologia do Banco de Portugal ao caso particular de Angola, no sentido de proceder à compilação de dados anuais e à produção das estatísticas de um grupo de empresas do setor de construção.

Para tal foi constituída uma amostra de 11 empresas portuguesas do setor da construção a operar em Angola, entre 2010 a 2012. Procedeu-se à imputação dos dados omissos através do método Cold-Deck e do método de imputação pela média para a elaboração do quadro de setor. Recorrendo à demonstração de resultados e ao balanço, os rácios económico-financeiros permitiram a análise da evolução económica e financeira e patrimonial do agregado do setor.

keywords

Angola; Bank of Portugal; cluster analysis; Cold- Deck method; Non-Financial Corporations; non-oil sector; sector framework

abstract

In the first part of the study we carried out an analysis of the similarity and dissimilarities measures between the 12 non-oil sector activity sectors (excluding the financial sector), which identified sectors clusters with a similar behavior and evolution between each other and dissimilar groups.

The cluster analysis allowed us to observe the existence of links and common features between the different sectors of activity, revealing sectoral dynamics that can be enhanced to introduce competitiveness and diversification of the Angolan economy.

Being the civil construction an organized sector with good growth prospects, it has been selected for the case study object and to this goal have been further developed the methodological and practical knowledge of the formulation of Non-Financial Corporations' statistics.

Thus, the case study intended to apply the methodology of the Bank of Portugal to the particular case of Angola, in order to proceed to the compilation of annual data and production statistics for a group of construction companies in the sector.

For this we have used a sample of 11 portuguese companies in the construction sector in Angola from 2010 to 2012. The procedure was the imputation of missing data through Cold-Deck method and the average allocation method for the preparation of the framework sector. Using the income statement and the balance sheet, the economic and financial ratios allowed the analysis of economic and financial developments and the balance of the sector aggregate.

Siglas e acromolós

ATA – Autoridade Tributária e Aduaneira

BACH – *Bank for Accounts of Companies Harmonised*

BAL – Balanço

BE – Boletim Estatístico

BC – Central de Balanços

BCE – Banco Central Europeu

BNA – Banco Nacional de Angola

BoP – Balança de Pagamentos

BdP – Banco de Portugal

CAE – Classificação da Atividade Económica

CRC – Central de Responsabilidades de Crédito

CMVMPSC – Custo das Mercadorias Vendidas e das Matérias-Primas e Subsidiárias Consumidas

DNLP - Logaritmo do Periodograma Normalizado

DDE – Departamento de Estatística do Banco de Portugal

DR – Demonstração de Resultados

ECCBSO – *European Committee of Central Balance-Sheet Data Offices*

EUA – Estados Unidos da América

FAO - *Food and Agriculture Organization*

FMI – Fundo Monetário Internacional

FUE – Ficheiro das Unidades Empresariais

IAHE – Inquérito Anual Harmonizado às Empresas

I&D – Investigação & Desenvolvimento

IDE – Investimento Direto Estrangeiro

IES – Informação Empresarial Simplificada

IACB – Inquérito Anual da Central de Balanços do Banco de Portugal

INE Angola – Instituto Nacional de Estatística de Angola

INE Portugal – Instituto Nacional de Estatística de Portugal

ITENF - Inquérito Trimestral às Empresas Não Financeiras

MPME – Micro, Pequenas e Médias Empresas

PALOP – Países de Língua Oficial Portuguesa

PIB – Produto Interno Bruto

PII – Posição de Investimento Internacional

PGC - Plano Geral de Contabilidade

POC – Plano Oficial de Contabilidade

QES – Quadros da Empresa e do Setor

QS – Quadros do Setor

SAE- Setores de Atividade Económica

SEE – Sistema Estatístico Europeu

SEBC – Sistema Europeu de Bancos Centrais

SGPS – Sociedades Gestoras de Participações Sociais

SIET – Sistema Integrado de Estatísticas de Títulos

SLQS – Série Longa dos Quadros do Setor

SNC – Sistema de Normalização Contabilística

SNF – Sociedades Não Financeiras

SPAI – Sistema de Partilha de Informação de Referência

SS – Segurança Social

USD – Dolar dos Estados Unidos da América

Índice

Siglas e acromolós	viii
Índice de Gráficos.....	xii
Índice de Quadros	xii
Índice de Figuras	xiii
Nota Introdutória.....	1
I. Análise de Conglomerados.....	5
1. Estatística de Similaridades dos Setores de Atividade Económica em Angola.....	6
1.1.Procedimentos e Técnicas de Análise Estatística de Similaridades	7
1.2.Estatística Descritiva	9
1.3.Estudo das Correlações entre os setores de atividade económica	16
1.4. Análise de Clusters.....	22
II. A Experiência do Banco de Portugal na Produção de Estatísticas sobre Sociedades Não Financeiras.....	33
2. Enquadramento Legal	33
3. Fontes de Informação e Controlo de Qualidade	34
3.1. Inquérito Trimestral às Empresas não Financeiras.....	34
3.2. Informação Empresarial Simplificada	34
3.3. Controlo de Qualidade.....	35
4. Compilação dos Dados Extrapolados.....	36
4.1. Tratamento da Informação Anual	36
4.2. Extrapolação de Dados Trimestrais.....	38
4.3. Conciliação dos Dados Anuais e Trimestrais.....	40
5. Produtos e Serviços Estatísticos da Central de Balanços do Banco de Portugal.....	42

5.1. Quadros de Empresa e Setor, Quadros de Setor e a Série Longa dos Quadros de Setor	42
5.2. Boletim Estatístico	43
5.3. Serviços	44
5.4. Portugal no Grupo BACH	44
III. Caso Prático: Aplicação da Produção de Estatísticas das Empresas Não Financeiras à realidade de Angola	47
6. Enquadramento Macroeconómico	48
6.1. Economia Internacional	48
6.2. Economia Angolana	49
7. Enquadramento Legal em Angola	51
8. Descrição do Problema em Estudo	52
9. Recolha da Informação	54
10. Tratamento da Informação Anual	56
10.1. Tratamento da Não Resposta	57
10.1.1. Imputação pelo Método da Média – Variante I	58
10.1.2. Imputação pelo Método Cold-Deck – Variante II	60
10.2. Seleção do Método	65
11. Análise Económico Financeira com Base no Quadro de Setor	66
11.1. Atividade e Rendibilidade	68
11.1.1. Volume de Negócios	68
11.1.2. Custos da Atividade Operacional	69
11.1.3. Rendibilidade dos Capitais Próprios	70
11.2. Situação Financeira	71
11.2.1. Estrutura Financeira	71
11.2.2. Custos Financeiros e Solvabilidade	72
11.3. Análise de Risco Económico-financeiro	73
Conclusões e Recomendações	75

Referências	83
ANEXOS.....	87

Índice de Gráficos

Gráfico 1. Gráfico de linhas sobre a evolução do PIB Mundial, entre 2008 e 2012 (%).....	48
Gráfico 2. Evolução do PIB, do PIB petrolífero e do PIB não petrolífero angolano (%)	50
Gráfico 3. Volume de Negócios do agregado das empresas em análise.....	68
Gráfico 4. Custos da atividade operacional	69
Gráfico 5. Rendibilidade dos Capitais Próprios.....	70
Gráfico 6. Grau de alavancagem combinada.....	73

Índice de Quadros

Quadro 1. Estatísticas descritivas dos setores de atividade económica	10
Quadro 2. Classificação dos Setores de Atividade em Três Classes	11
Quadro 3. Avaliação de <i>Outliers</i>	14
Quadro 4. Interpretação do Coeficiente de Correlação de Pearson	17
Quadro 5 - Matriz de correlações entre os setores de atividade económica em estudo.....	18
Quadro 6. Coeficiente de correlação cofenética	26
Quadro 7. Sumário dos clusters/grupos obtidos	29
Quadro 8. Pressupostos Comuns para a Variante I e Variante II	57
Quadro 9. Pressupostos Específicos da Variante I e Variante II.....	57
Quadro 10. Balanço da EMP1 (Em Milhões de Kwanzas).....	62
Quadro 11. Balanço das EMP2 e EMP8 (Em Milhões de Kwanzas)	63
Quadro 12. Total de ativo 2010-2012 (Em Milhões de Kwanzas).....	64
Quadro 13. Indicadores sínteses da Variante I e Variante II, para os anos de 2010, 2011 e 2012 ..	65
Quadro 14. Indicadores económico-financeiros.....	67
Quadro 15. Autonomia Financeira: Distribuição por quartis e média agregada	71
Quadro 16. Caracterização do agregado do Quadro do Setor	87
Quadro 17. Rácios Setoriais Europeus (a partir de 2006) - Base de Dados BACH para França e Portugal	88

Índice de Figuras

Figura 1. A taxa de distribuição das empresas por setor de atividade	8
Figura 2. <i>Boxplots</i> da Classe de CAE de Baixa Percentagem	12
Figura 3. <i>Boxplots</i> da Classe de CAE de Média Percentagem	13
Figura 4. <i>Boxplots</i> da Classe de CAE de Alta Percentagem.....	15
Figura 5. Diagrama de <i>Scatterplots</i>	19
Figura 6. Dissimilaridade baseada na correlação	28
Figura 7. Dissimilaridade baseada no periodograma simples.....	28
Figura 8. Imputação pela Média – Variante I	59
Figura 9. Dissimilaridade baseada na autocorreção parcial (critério average linkage e critério centroide)	89
Figura 10. Dissimilaridade baseada na autocorreção parcial (critério average linkage e critério median linkage).....	89
Figura 11. Dissimilaridade baseada no periodograma integrado (critério average linkage e critério McQuitty)	90

Nota Introdutória

A escolha do tema “Análise de Conglomerados e Estatísticas das Sociedades Não Financeiras”, aplicado ao caso particular de Angola, prende-se, em primeiro lugar, com o objetivo de identificar grupos ou *clusters* de Setores de Atividade Económica (SAE), mediante a análise das características de similaridade e dissimilaridades do conjunto dos setores que integram o setor não petrolífero (com exceção do setor financeiro). Em segundo lugar, dado o interesse atual de Angola na produção de estatísticas das Sociedades Não Financeiras (SNF), pretende-se aplicar a metodologia utilizada pelo Banco de Portugal (BdP) na produção dos Quadros de Setor para o caso específico do setor da construção. A escolha do setor da construção prendeu-se com a sua organização, dimensão e importância para a economia angolana, no que respeita ao seu contributo para o Produto Interno Bruto (PIB) do setor não petrolífero ao ocupar a terceira posição em 2012 e consequentemente para o PIB angolano. (Contas Nacionais 2007-2012; outubro 2014).

O estudo parte assim, de uma abordagem macro, no que se refere a análise de informação agregada dos setores de atividade económica do setor não petrolífero, isto é, (número total de empresas ativas que compõem cada setor), no período de 2003 a 2014, para uma abordagem micro na qual serão analisadas as estatísticas das SNF do setor da construção, no período de 2010 a 2012.

Para tal, procedeu-se, numa primeira etapa, ao estudo das medidas de similaridade e dissimilaridade do universo dos 12 setores económicos (com base na Classificação das Atividades Económicas de Angola (CAE-Rev.2)) que incorporam as empresas ativas que atuaram em Angola, durante um período de 12 anos (2003-2014). Foi realizada a análise exploratória dos dados em estudo, com recurso a medidas de localização, dispersão e forma da distribuição dos dados, no sentido de descrever o comportamento de cada setor. Com vista à identificação das correlações entre os setores de atividade económica foi analisado o coeficiente de correlação de Pearson.

A análise das medidas de similaridade e dissimilaridades do comportamento dos setores de atividade recorreu ao método não supervisionado (que pode servir também para extrair características dos dados e desenvolver hipóteses a respeito da sua natureza), não paramétrico designado por análise de conglomerados (*Cluster Analysis*) com recurso ao “*TSclust package*” do *software* estatístico *R*. Esta ferramenta permite efetuar o conglomerado hierárquico de variáveis, geração de resultados e respetiva análise. Para o conglomerado hierárquico dos setores de atividade foram utilizadas as medidas de similaridades livres de modelos nomeadamente: a

correlação, a autocorrelação simples e parcial, bem com o periodograma simples e integrado, associados a sete critérios de agregação hierárquica (*“complete linkage”, “single linkage”, “average linkage”, “median linkage”, “centroide linkage”, “ward linkage”* e *“Mc Quitty”*). Através da análise do coeficiente de correlação cofenético identificaram-se os critérios e as medidas de similaridades mais ajustados.

A identificação de características de homogeneidade e de heterogeneidade entre os setores de atividade económica permitiu estudar o comportamento e a evolução dos diferentes setores. Este conhecimento útil é uma mais-valia para Angola, numa altura em que o Executivo está a delinear políticas orientadas para uma maior diversificação da economia e uma maior autonomia da mesma face ao setor petrolífero.

A segunda etapa do estudo integra a experiência do BdP na produção de estatísticas das SNF, em termos das normas que regulam a produção das referidas estatísticas, as fontes de informação contabilística e as técnicas de controlo a qualidade utilizadas pelo BdP, bem como as técnicas de imputação, extrapolação e conciliação dos dados contabilísticos.

A última etapa corresponde ao caso prático da produção das estatísticas das SNF aplicada à realidade de Angola e teve como objetivo o estudo e a aplicação da metodologia do BdP na produção das estatísticas das SNF. Depois de realizado o enquadramento macroeconómico da economia internacional e da economia angolana relativo ao ano de 2012, foi estudada uma amostra de 11 empresas portuguesas do setor da construção a operar em Angola, tendo como período de observação os anos 2010, 2011 e 2012. Através de métodos de imputação de dados omissos (imputação pela média e Cold-Deck), procedeu-se ao tratamento da informação anual da não resposta. Obtidos os mapas da demonstração de resultados e do balanço, para o caso específico do setor da construção, foi analisada a evolução económica e financeira e patrimonial do agregado através do quadro de setor.

A par dos dados micro do setor da construção, no que se refere à compilação, controlo de qualidade da informação, tratamento de dados anuais, extrapolação de dados trimestrais, conciliação de dados trimestrais com os anuais e produção de estatísticas das SNF, a metodologia do BdP, constituiu o objeto de estudo do caso prático, na medida em que foram utilizadas as técnicas e os procedimentos metodológicos que podem possíveis de ser adaptados à realidade de Angola.

A informação produzida pelo BdP sobre as SNF, na ótica contabilística, conjugada com outras fontes de informação, permite caracterizar e acompanhar a situação económica e financeira, a

evolução patrimonial e a informação sobre um conjunto de variáveis que caracterizam a situação das mesmas, tais como a atividade económica, o emprego, os resultados, rendibilidades, passivos e ativos financeiros.

A produção de estatísticas sobre as SNF, não só de dados micro, como também macro, é uma matéria que tem vindo a registar um interesse significativo por parte dos Países de Língua Oficial Portuguesa (PALOP) e dos países da África Austral, em particular de Angola. A compilação deste tipo de estatísticas constitui uma importante ferramenta de análise da situação de competitividade da economia nacional.

Acredita-se que em Angola a compilação e a produção de informação estatística sobre o comportamento dos setores económicos e a sua estabilidade financeira potenciam o investimento direto estrangeiro e a implementação de políticas de investimento público visando a diversificação da economia. Atualmente, a publicação das estatísticas sobre a evolução das empresas não financeiras é realizada pelo Instituto Nacional de Estatística (INE) de Angola, através do Inquérito Anual Harmonizado às Empresas (IAHE) que recolhe dados económico-financeiros numa base anual, nomeadamente a Demonstração de Resultados.

O IAHE incorpora informação detalhada sobre os investimentos realizados pelas empresas e os serviços de terceiros, mas não contempla informação do balanço, fluxo de caixa, nem as notas às contas das empresas. A informação recolhida pelo INE Angola é reportada de acordo com o Plano Geral de Contabilidade e obedece às normas contabilísticas internacionais.

Do ponto de vista da análise de conglomerados, os resultados revelaram a formação de *clusters* de dimensão variada, mas cuja constituição confirma a existência de sinergias e complementaridades entre os *clusters* obtidos. Acredita-se que estes resultados podem constituir o suporte para a implementação de medidas de política que visem o estímulo à competitividade e à diversificação da economia angolana.

Uma vez concretizados os objetivos da investigação, e não obstante as limitações e constrangimentos associados ao trabalho do investigador, o último ponto do trabalho apresenta as considerações finais sobre a análise de conglomerados dos setores de atividade económica, bem como as conclusões do caso prático desenvolvido no âmbito do estágio realizado no BdP. No sentido de potenciar o aprofundamento do tema, são apresentadas recomendações para futuras investigações e abordagens à análise de *clusters* dos setores de atividade em Angola e à produção das estatísticas das SNF.

I. Análise de Conglomerados

A diversificação da economia é uma prioridade para a República de Angola. Por um lado, para a substituição paulatina das importações e aumento da produção interna de forma a satisfazer a procura doméstica e, por outro, para melhorar a estrutura das exportações e assegurar a redução da dependência económica e financeira do petróleo, uma vez que mais de 90% das receitas de exportação são provenientes do setor petrolífero. Assim, com a forte queda do preço em dólar dos Estados Unidos da América (USD) do barril do petróleo, em cerca de 60% desde 2009 no mercado internacional, tornou-se imperativo acelerar o processo de diversificação da economia angolana.

Através da análise de conglomerados (*clusters*) pretendeu-se constituir grupos ou *clusters* de setores de atividade económica com comportamento e características semelhantes entre si e dissemelhantes face aos restantes grupos. Combinada com uma análise exploratória, esta abordagem teve como objetivo conhecer a evolução de cada setor no período em análise e identificar relações entre os mesmos, de forma a melhor perceber as dinâmicas setoriais. Este conhecimento é pertinente quando se planeiam incentivos para os diferentes setores, nomeadamente, incentivos económicos que apoiem o processo de diversificação da economia nacional, o aumento da competitividade e a redução da dependência da economia nacional face ao setor petrolífero.

A análise de *clusters* inclui um conjunto de procedimentos de estatística multivariada usados para classificar um conjunto de entidades (objetos ou indivíduos), para os quais é conhecida informação detalhada, em grupos relativamente homogêneos (Reis, 2001). Para um conjunto de n entidades, para os quais existe informação sob a forma de p variáveis, o método de análise de *clusters* procede ao conglomerado dos indivíduos, em função da informação existente, de tal modo que os indivíduos pertencentes a um mesmo grupo sejam tão semelhantes quanto possível e sempre mais semelhantes aos elementos do mesmo grupo do que a elementos dos restantes grupos (Reis, 2001). Um *cluster* pode ser definido como sendo um conglomerado composto por um número não fixo de objetos (exemplos) similares, de acordo com uma medida de similaridade (Everitt, Landau & Leese, 2001).

Com base na observação das semelhanças ou dissemelhanças entre os objetos, e sem definir previamente critérios de inclusão de qualquer conglomerado, o método da análise de *clusters* (também chamado de classificação ou taxonomia) tem sido utilizado em tarefas de exploração de dados e identificação de padrões (Reis, 2001).

1. Estatística de Similaridades dos Setores de Atividade Económica em Angola

A técnica do reconhecimento de padrões desenvolveu-se significativamente como um campo de estudo na década de 1960 e tornou-se uma técnica interdisciplinar que abrangia as áreas da estatística, engenharia, inteligência artificial, ciência da computação, psicologia e ciências sociais em geral. O desenvolvimento dos sistemas baseados no conhecimento, na década de 1970, e o estudo das redes neurais, na década de 1980, permitiram o aperfeiçoamento das técnicas de reconhecimento de padrões, sendo hoje uma componente fundamental da maior parte dos sistemas de tomada de decisão.

Os padrões são representados por vetores de atributos obtidos através de amostras cujos dados podem estar total ou parcialmente rotulados, ou não rotulados. Dependendo da quantidade de informação disponível na base de dados e dos objetivos que se pretendem atingir, podem ser identificados dois tipos de reconhecimento de padrões: o método supervisionado (que incluiu, por exemplo, a análise fatorial e a análise discriminante) e o não supervisionado (referida também na literatura simplesmente como classificação ou análise de *clusters*) (Webb, 2002).

Enquanto que no reconhecimento supervisionado os dados encontram-se categorizados e agrupados, no reconhecimento não supervisionado os dados não estão categorizados e, por isso, procura-se identificar grupos com base nas características que simultaneamente os distinguem dos outros e aproximam os elementos de cada grupo. É exemplo deste tipo de reconhecimento a análise de *clusters* (Webb, 2002).

Com base na análise das características de similaridade e dissimilaridade dos setores de atividade económica em Angola do setor não petrolífero (com a exclusão do setor financeiro), e tendo em conta a distribuição das empresas ativas em Angola por setor de atividade económica, sujeito à Classificação das Atividades Económicas de Angola (CAE- Rev.2), foram identificados conglomerados, ou *clusters*, de setores com características semelhantes e comportamento homogéneo.

Para tal, procedeu-se numa primeira fase à análise exploratória do comportamento e evolução dos diferentes setores de atividade, no sentido de identificar características que possibilitem a caracterização de grupos de setores. Através dos métodos exploratórios, foram geradas suposições sobre o conglomerado dos setores de atividade económica com características semelhantes entre si e, simultaneamente, dissemelhantes em relação aos restantes conglomerados.

1.1. Procedimentos e Técnicas de Análise Estatística de Similaridades

Com vista a análise estatística de similaridades dos setores de atividade económica que compõem o setor não petrolífero em Angola, procedeu-se à análise exploratória dos dados. A análise exploratória, ou estatística exploratória, foi uma técnica introduzida por John Wilder Tukey no livro *Exploratory Data Analysis* (1977), sendo bastante utilizada para a análise de dissimilaridades.

Assim, a análise exploratória de dados teve como base a estatística descritiva, tendo sido utilizadas medidas de localização e tendência central (como a média e a mediana), quartis e extremos (inferior e superior), medidas de dispersão como o desvio padrão, a amplitude interquartil, o intervalo de variação e o coeficiente de variação, e medidas de assimetria, bem como descrições gráficas dos dados através de caixas de bigodes (*boxplots*) para visualização da estrutura dos dados. (Pinto,2009).

A análise exploratória dos dados incidiu sobre as medidas de estatística descritiva da amostra (N=12). A base de dados foi fornecida pelo INE Angola e continha a informação relativa à distribuição das empresas pelos 14 setores de atividade económica observada durante um período de doze anos, de 2003 a 2014. Do conjunto inicial dos 14 setores de atividade económica, foram excluídos dois setores: o financeiro e o da indústria extrativa.

O setor petrolífero gera mais de 90% das receitas de exportação em Angola que suporta o desenvolvimento da economia angolana, tornando a mas vulnerável aos choques externos. Neste contexto torna-se imperioso inverter esta situação, razão pela qual se pretende estudar as potencialidades do setor não petrolífero. Tendo em consideração que os país depende das importação de bens e serviços para satisfação da sua procura interna, serão, portanto, objeto de estudo as variáveis referentes aos restantes 12 setores de atividade económica que compõem o setor não petrolífero, com vista a identificação de setores que poderão contribuir para a diversificação da economia angolana.

Inicialmente os dados fornecidos correspondiam à contagem de empresas distribuídas por cada setor. Para evitar a disparidade dos valores dessas contagens, optou-se por trabalhar em termos relativos (0-100%).

Os dados foram apresentados como taxas de distribuição das empresas ativas por setores de atividade económica, devido à diversidade de valores que as variáveis assumiam. Assim, a taxa de distribuição é dada pela seguinte equação matemática:

$$T_{ex} \equiv \frac{n_x}{N} \times 100\%$$

onde:

T_{ex} : Taxa de distribuição das empresas ativas por setor de atividade económica

e : Empresas

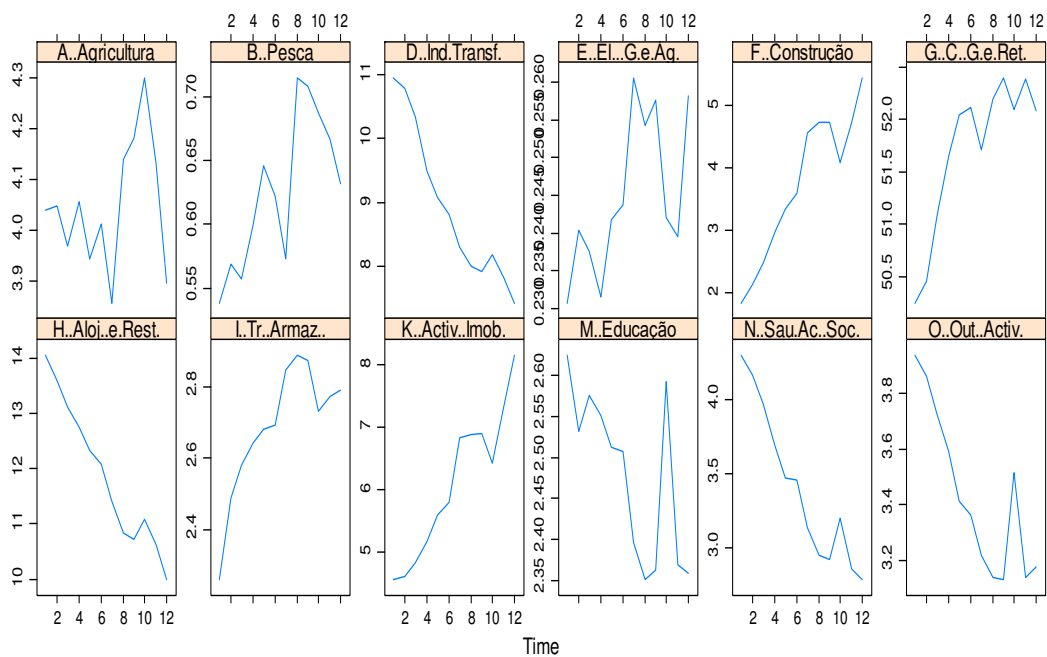
x : setor de atividade económica

n_x : Número de empresas do setor de atividade económica x

N : Total de empresas ativas

Assim, começou-se por analisar o gráfico *xyplot*, gerado pelo *package Lattice* do *software R*, que espelha o comportamento de evolução das variáveis individuais (setor de atividade económica) em estudo (Figura 1), com vista à formulação de algumas suposições que poderão ser confirmadas ao longo do estudo através da análise de clusters.

Figura 1. A taxa de distribuição das empresas por setor de atividade



Na Figura 1 verifica-se que de 2003 a 2014 a distribuição das empresas por setor de atividade foi diferente e variada, não permitindo nesta fase formular conglomerados de setores de atividade com base nas características de similaridade e dissimilaridade. A análise da Figura 1

sugere que os setores da indústria transformadora, do alojamento e restauração, da saúde e ação social e o de outras atividades possam ser agrupados no mesmo *cluster*, pois apresentam um comportamento decrescente e semelhante.

Os setores da construção, do comércio, grosso e retalho, dos transportes e armazenagem, e da atividade imobiliária apresentam também um comportamento semelhante com tendência crescente, o que pressupõe que poderão constituir o mesmo conglomerado.

Num outro possível *cluster*, poderão ser integrados os setores da agricultura, das pescas e o da eletricidade, gás e água, os quais aparentam possuir um comportamento semelhante.

A educação, por sua vez, registou um comportamento distinto comparativamente aos restantes setores, o que não permitiu identificar um conglomerado onde o mesmo pudesse ser integrado.

1.2. Estatística Descritiva

A análise da estatística descritiva permitiu observar as medidas de localização central, como a média e a mediana (Q2), de localização não central nomeadamente o primeiro e terceiro quartis (Q1 e Q3), os extremos valor mínimo e máximo, as medidas de dispersão, tais como o desvio padrão, coeficiente de variação e intervalo de variação para a taxa de empresas ativas por setor de atividade económica de acordo com a CAE-Rev.2, ao longo do período de 2003 a 2014 (Quadro 1).

Quadro 1. Estatísticas descritivas da taxa de empresas ativas por setor de atividade económica

SAE	Mínimo	1º Quartil	Mediana (2º Quartil)	Média	Desvio Padrão	Coefficiente de Variação	3º Quartil	Máximo	Intervalo Variação
A - Agricultura	3,86	3,96	4,04	4,05	0,126	3,11	4,13	4,30	0,44
B - Pesca	0,54	0,57	0,63	0,63	0,063	10,00	0,67	0,72	0,18
D - Ind.Transf.	7,43	7,98	8,55	8,92	1,213	13,60	9,70	10,95	3,52
E - El...G.e.Ág.	0,23	0,24	0,24	0,24	0,009	3,75	0,26	0,26	0,03
F - Construção	1,84	2,84	3,84	3,72	1,178	31,67	4,73	5,45	3,61
G - C. G.e.Ret.	50,25	51,51	52,07	51,71	0,727	1,41	52,13	52,39	2,14
H - Hot. Aloj..e.Rest.	9,99	10,80	11,75	11,89	1,308	11,00	12,85	14,07	4,08
I - Tr..Armaz..	2,26	2,63	2,71	2,69	0,180	6,69	2,81	2,89	0,63
K - Activ..Imob.	4,55	5,08	6,11	6,09	1,722	28,28	6,88	8,16	3,61
M - Educação	2,35	2,37	2,51	2,48	0,103	4,15	2,56	2,63	0,27
N - Sau.Ac..Soc.	2,79	2,94	3,33	3,41	0,525	15,40	3,76	4,30	1,52
O - Out..Activ.	3,13	3,17	3,39	3,43	0,291	8,48	3,62	3,94	0,81

De acordo com as estatísticas descritivas da taxa de empresas ativas por setores de atividade económica entre 2003 e 2014, observa-se pelo Quadro 1 que foi o setor da eletricidade, gás e água o que apresentou valores extremos da distribuição muito próximos (Min.=0,23; Max.=0,26). O valor do desvio padrão face à média (M=0,24; DP=0,01) refletiu uma reduzida dispersão dos dados, ou seja, os valores aproximaram-se da média. Além disso o valor da média foi igual ao da mediana (Q2=0,24) o que sugere uma distribuição simétrica, logo os quartis encontram-se equidistantes da mediana.

O setor das pescas, apresentou um comportamento semelhante ao da eletricidade, gás e água no que concerne aos extremos de 0,54 (Min.) a 0,72 (Max.). Também aqui o valor da média coincidiu com o valor da mediana (Q2=0,63) e o desvio padrão foi baixo (M=0,63; DP=0,06) o que demonstra uma reduzida variabilidade dos dados neste setor.

A construção apresentou o maior coeficiente de variação (Cv=31,67), logo apresenta uma maior dispersão dos dados comparativamente aos setores em análise, bem como o setor de saúde e ação social que apresentou o terceiro maior coeficiente de variação (Cv=15,40) comparativamente aos restantes setores.

Apesar de o setor da atividade imobiliária registar o maior desvio padrão ($M=6,09$; $DP=1,72$) o que revela que apresenta uma grande variabilidade da distribuição na taxa de empresas ativas em torno da média no setor. Comparativamente aos restantes setores, apresenta o segundo maior coeficiente de variação ($Cv=28,28$), que indica uma grande dispersão dos valores em torno das médias, de distribuição de frequências dos restantes setores.

O setor do comércio, grosso e retalho apesar de apresentar a maior média ($M=51,71$) comparativamente aos demais setores. A mediana ($Q2=52,07$) situou-se próxima do valor máximo ($Max.=52,39$) o que significa que 50% das observações apresentaram valores muito próximos e próximos do máximo, foi o setor que registou o menor coeficiente de variação ($Cv=1,41$) comparativamente aos restantes setores, refletindo assim um maior grau de concentração das observações em torno da média.

O setor da educação e o dos transportes e armazenagem apresentaram valores médios semelhantes ($M=2,48$; $DP=0,10$ e $M=2,69$; $DP=0,18$ respetivamente). Entretanto no setor dos transportes e armazenagem o coeficiente de variação ($Cv=6,69$), foi superior ao observado no setor da educação ($Cv=4,15$), revelando assim uma menor variabilidade dos dados na educação.

Visando a melhoraria da visualização gráfica dos dados, foram definidas três classes de setores de atividade económica com base na taxa média de empresas por sector de atividade: Baixa percentagem, Média e Alta percentagem. Tal classificação, foi efetuada com base nos intervalos de $[0,00; 1,00[$, $[1,00; 5,00[$ e $[5; +\infty]$ respetivamente, conforme sistematiza o Quadro 2.

Quadro 2. Classificação dos setores de atividade económica em Três Classes

Classes	Setores de atividade	Intervalo da Classe
Baixa Percentagem	2	$[0,00; 1,00[$
Média Percentagem	6	$[1,00; 5,00[$
Alta Percentagem	4	$[5,00; +\infty[$

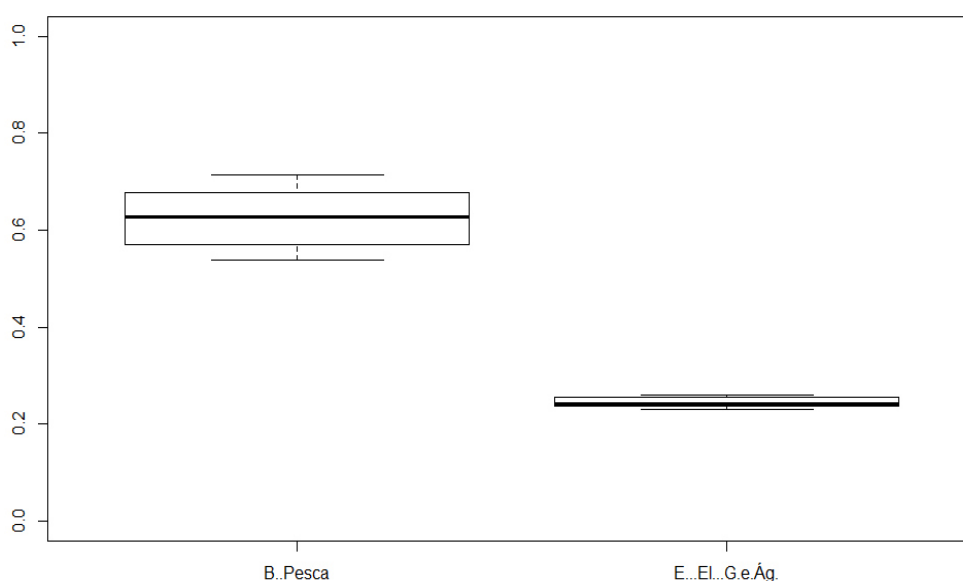
Os casos em que a média se situou entre 0,00 e 1,00 os setores de atividade económica foram considerados de Baixa Percentagem e contabilizaram-se dois setores de atividade económica nesta classe. Quando a média se situou entre 1,00 e 5,00 os setores de atividade económica foram considerados de Média Percentagem e nesta classe incluíram-se seis setores de atividade. A Alta Percentagem correspondeu aos setores com média igual ou superior a 5,00 e integrou quatro setores.

Recorreu-se à representação dos dados através de *boxplots* os quais fornecem informação sobre a localização, dispersão e *outliers*, correspondentes a observações discrepantes, isto é, que fogem ao padrão dos dados.

Nas Figuras 2, 3 e 4 são apresentadas caixas de bigodes comparativas para os setores de atividades que integram cada classe.

Na Figura 2 estão reportados os *boxplots* relativos aos setores da pesca e da eletricidade, gás e água.

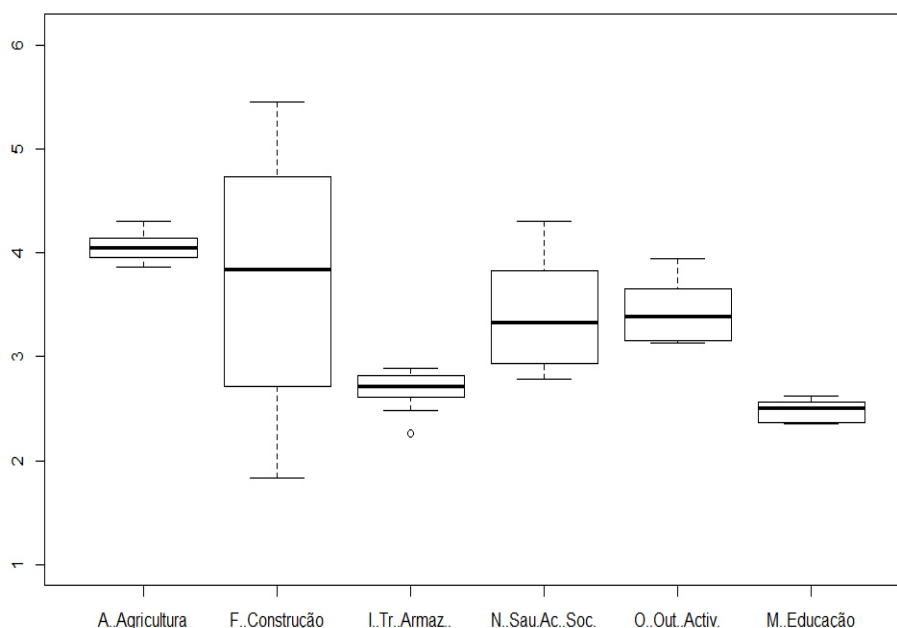
Figura 2. *Boxplots* da Classe de setores de atividade de Baixa Percentagem



Pela Figura 2 verificou-se que as distribuições dos dois setores são semelhantes. Tendo em conta a informação apresentada no Quadro 1, constatou-se que os setores das pescas e da eletricidade, gás e água apresentaram uma distribuição simétrica, pois a sua média é igual à mediana. Os dois setores apresentaram coeficientes de variação de 10,00% e 3,75% (Quadro 1), respetivamente, o que indica que os dois setores apresentaram baixa dispersão de dados, abaixo dos 15,00%, estão próximos da média e os dados são homogéneos.

Na Figura 3 estão representados os *boxplots* da Classe de setores de atividade de Média Percentagem. A sua análise reforçou a ideia de que as distribuições dos setores de atividade são distintas entre si e evidencia diferenças entre os valores da mediana e a amplitude entre o máximo e o mínimo.

Figura 3. Boxplots da Classe de setores de atividade de Média Percentagem



Com base na informação do Quadro 1, constatou-se que nos setores da agricultura, saúde e ação social e o de outras atividades, a média é superior à mediana, logo pode concluir-se que a sua distribuição é assimétrica positiva, pois os valores aproximaram-se do 1º quartil, verificando-se a concentração de dados em valores mais baixos, a média sofrerá influência da cauda à direita deslocando-se em sua direção, logo haverá mais observações abaixo da média do que acima dela.

Nos restantes setores da classe de média percentagem, a média é inferior à mediana, logo a distribuição é assimétrica negativa, pois aproxima-se do 3º quartil, verificando a concentração em valores mais altos e a média sofre a influência da cauda à esquerda e desloca-se em sua direção, logo haverá mais observações acima da média do que abaixo dela.

Nos setores da construção e da saúde e ação social o coeficiente de variação foi de 31,67 % e 15,40%, respetivamente, o que indica que o setor da construção apresenta alta dispersão os dados são heterogêneos ou seja, as observações estão afastadas do valor médio, enquanto a saúde e ação social apresenta média dispersão, valores superiores a 15,00%. e inferior a 30,00%.

Os setores da agricultura, transporte e armazenagem, outras atividades e educação apresentaram coeficientes de variação de 3,11%; 6,69%; 8,48%; 4,15% respetivamente (Quadro 1), os setores apresentaram baixa dispersão, inferior a 15,00%, logo os valores são homogêneos e estão mais próximas do valor médio comparativamente aos restantes setores da classe.

No Quadro 3 analisa-se a existência de *outliers*, que é confirmada no setor dos transportes e armazenagem uma vez que o valor de 2,26 se situou fora dos extremos do *boxplot* [2,36; 3,07]. É uma observação discrepante localizada próxima do extremo inferior do *boxplot* tendo sido considerado estatisticamente um *outlier* não severo, se a este aspeto se associa a avaliação do ponto de vista económico, percebe-se que o ano de 2003 se refere ao início das observações e esta variável apresenta um comportamento crescente ao longo do período em análise, razão pela qual se decidiu pela sua não exclusão do conjunto de dados (Figura 3).

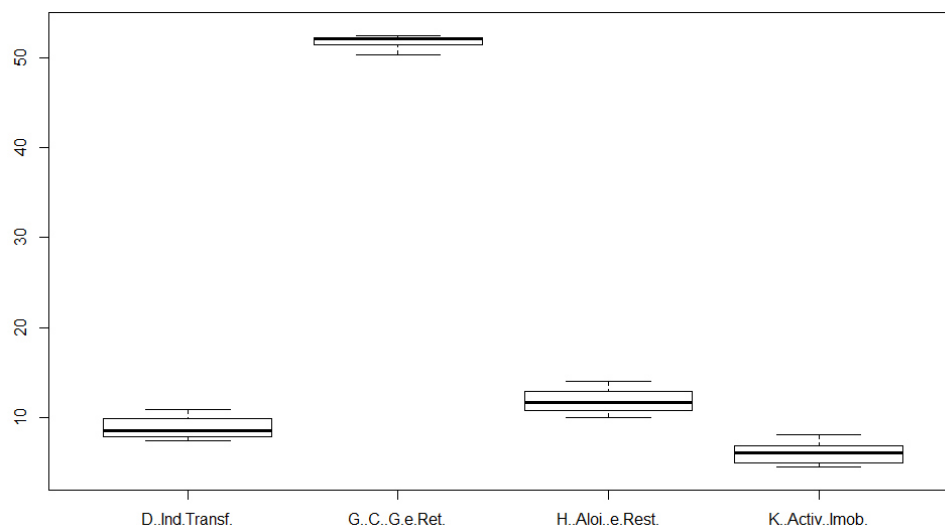
Quadro 3. Avaliação de *Outliers*

Setor de atividade económico	Q1	Q3	IIQ=Q3-Q1	Limite Inferior Q1-1,5*IIQ	Limite Superior Q1+1,5*IIQ
A - Agric.	3,963	4,134	0,171	3,707	4,391
B - Pesc.	0,572	0,672	0,101	0,421	0,823
D - Ind.T.	7,977	9,697	1,720	5,397	12,277
E - El.G.e.Ág.	0,239	0,255	0,016	0,215	0,279
F - Const.	2,838	4,733	1,895	-0,004	7,576
G - C..G.e.R.	51,510	52,130	0,620	50,580	53,060
H - Aloj..e.R.	10,803	12,853	2,050	7,728	15,928
I - Tr..Armaz..	2,627	2,805	0,178	2,360	3,072
K - Activ..Imob.	5,080	6,883	1,803	2,376	9,588
M - Educação	2,368	2,557	0,189	2,085	2,841
N - Sau.Ac..Soc.	2,944	3,760	0,816	1,720	4,984
O - Out..Activ.	3,167	3,622	0,455	2,485	4,305

Nota: IIQ=Intervalo Interquartil

Pela análise da Figura 4, verifica-se que a distribuição do setor comércio, grosso e retalho é muito distinta dos restantes setores desta classe de alta percentagem.

Figura 4. Boxplots da Classe de setores de atividade de Alta Percentagem



Os setores da indústria transformadora e o do alojamento e restauração apresentaram uma distribuição assimétrica positiva, visto que a média é superior à mediana, pois a mediana aproxima-se do 1º quartil. Por outro lado, para os setores do comércio a grosso e a retalho e o da atividade imobiliária a média é inferior à mediana, logo a distribuição é assimétrica negativa, pois a mediana aproxima-se do 3º quartil.

Os quatro setores apresentaram coeficientes de variação de 13,60%; 1,41%; 11,00%; 28,28% (Quadro 1), o setor de indústria transformadora, o comércio a grosso e a retalho e o alojamento e alimentação apresentaram baixa dispersão, enquanto o setor de atividade imobiliária apresentou média dispersão.

Após a análise descritiva dos setores de atividade económica, seguiu-se a análise de correlação entre os mesmos, com vista a determinar a sua direção e grau de intensidade da mesma, conforme se apresenta no ponto seguinte.

1.3. Estudo das Correlações entre os setores de atividade económica

A teoria da análise de correlação teve início na segunda metade do século XIX. Francis Galton (1822-1911) foi quem usou pela primeira vez os termos correlação e regressão. Publicou em 1869 o livro *Hereditary Genius*, sobre a teoria da regressão (Schultz & Schultz, 1992).

O termo “correlação” significa “co-relacionamento” e indica até que ponto os valores de uma variável estão relacionados com os de outra, através de uma medida que “sintetize” o grau de relacionamento entre as variáveis. É importante promover-se a análise de correlação, com vista a determinar o grau de relacionamento entre os setores de atividade económica e, com esta informação, descrever o que se espera para um setor com base no conhecimento de outro.

Pode-se inferir sobre um setor com base no outro. De salientar que não existe distinção entre a variável explicativa e a variável resposta, ou seja, o grau de variação conjunta entre X e Y é igual ao grau de variação entre X e Y (Mukaka, 2012).

Nesta parte do estudo aplica-se a análise do coeficiente de correlação linear de Pearson, com o objetivo de estudar o sinal da correlação (positiva ou negativa) e o grau de associação linear entre o número de empresas de cada de setor de atividade económico, medidos na métrica de razão.

Tratando-se de variáveis de razão, foi utilizado o coeficiente de correlação de Pearson adequado a relações lineares entre duas variáveis quantitativas X, Y também chamado ρ de Pearson que é dado pela seguinte expressão:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{cov(X, Y)}{\sqrt{var(X)var(Y)}}$$

onde x_1, x_2, \dots, x_n e y_1, y_2, \dots, y_n são as observações das séries temporais X e Y . Note-se que $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ e $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ são as médias aritméticas das séries temporais X e Y , respetivamente.

O coeficiente de correlação de Pearson é uma medida do grau de relação linear entre duas variáveis quantitativas. Este varia entre os valores -1 e 1 . O valor 0 significa que não há relação linear, enquanto os valores -1 e 1 indicam uma relação linear perfeita. Quanto mais próximo este coeficiente estiver de -1 ou 1 , mais forte é a associação linear entre as duas variáveis (Mukaka, 2012). Este coeficiente pode ser interpretado da seguinte forma:

$\rho = -1$: Correlação negativa perfeita entre as duas variáveis, ou seja, quando uma variável aumenta a outra diminui de forma inversamente proporcional;

$\rho = 0$: Não existe correlação linear entre as variáveis;

$\rho = 1$: Correlação perfeita positiva entre as duas variáveis, ou seja, quando uma variável aumenta a outra também aumenta de forma diretamente proporcional.

A intensidade da correlação pode variar em valor absoluto entre 0 (nula) e 0,3 (desprezível) a 0,9 (muito forte). A interpretação do valor da correlação pode ser feita com base nos intervalos de valores propostos por Mukaka (2012) conforme sistematiza o Quadro 4.

Quadro 4. Interpretação do Coeficiente de Correlação de Pearson

Intervalo de confiança				
Limite Inferior	Limite Superior	Sentido		Correlação
	0,9			Muito forte
0,7	0,9	Negativa	Positiva	Forte
0,5	0,7	Negativa	Positiva	Moderada
0,3	0,5	Negativa	Positiva	Fraca
0	0,3	Negativa	Positiva	Desprezível

Fonte: (Mukaka, 2012)

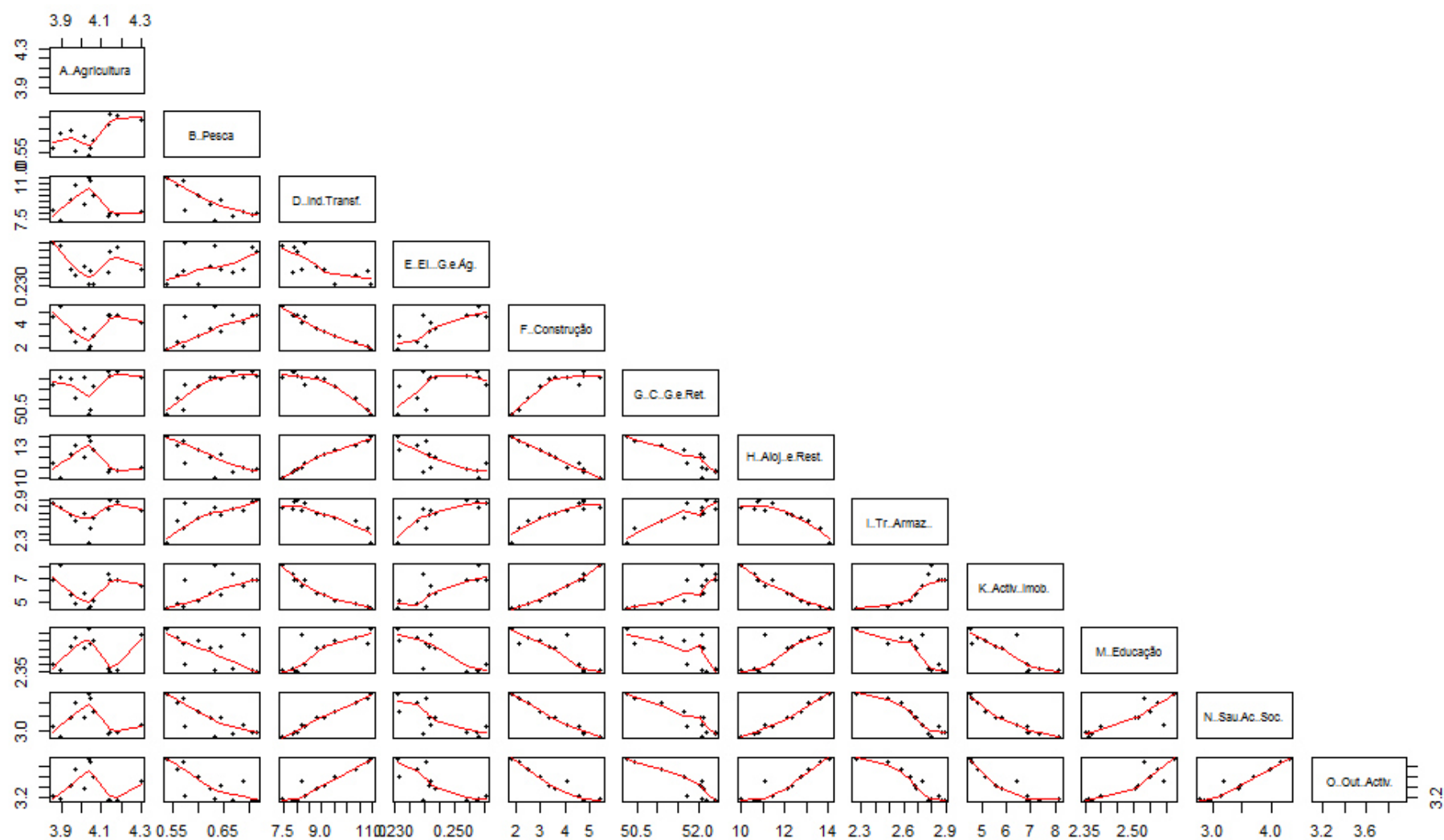
O estudo do coeficiente de correlação entre os setores de atividade permitiu identificar dois grupos, os que se influenciam de forma inversa (negativa) e os que se influenciam de forma direta (positiva). O Quadro 5 representa a matriz de correlações (matriz simétrica) com os valores do coeficiente de Pearson que permite avaliar as correlações obtidas.

Para completar a análise da correlação entre os setores das atividades económicas, apresenta-se o Diagrama de *Scatterplots*, para visualizar graficamente a relação/associação existente entre os distintos setores de atividade (Figura 5). O Diagrama de *Scatterplots* (diagrama de dispersão cartesiano, ou seja, diagramas X, Y) é considerado o melhor método para examinar os dados no que se refere à ocorrência de tendências lineares e não lineares dos conglomerados de duas ou mais variáveis, mudanças de dispersão de uma variável em relação à outra e verificar a ocorrência dos valores discrepantes. São consideradas relações lineares, quando os pontos no gráfico (X, Y) se aproximam do padrão de uma reta linear. As relações lineares são diretas positivas quando as duas variáveis crescem no mesmo sentido e inversas negativas quando as variáveis crescem em sentidos opostos.

Quadro 5 - Matriz de correlações entre os setores de atividade económica

	Agric.	Pesca	Ind.Tran.	El.G.e.Ág.	Const.	C.G.e.Ret.	H.Al.e.R.	I.Tr.Arm.	K.Acti.Im.	M..Educ	N.S. Ac.S	Out.Activ.
Agricultura	1											
Pesca	0,617	1										
Ind.Transf.	-0,169	-0,776	1									
Elet. G. Ág.	-0,189	0,442	-0,715	1								
Construção	0,080	0,714	-0,984	0,796	1							
C.G.e.Ret.	0,242	0,835	-0,920	0,528	0,853	1						
Aloj..e.Rest.	-0,189	-0,770	0,985	-0,736	-0,988	-0,873	1					
Tr..Armaz..	0,111	0,746	-0,904	0,776	0,900	0,890	-0,887	1				
Activ..Imob.	0,049	0,648	-0,956	0,758	0,984	0,783	-0,976	0,815	1			
Educação	0,094	-0,567	0,786	-0,806	-0,858	-0,655	0,810	-0,805	-0,843	1		
Sau.Ac..Soc.	-0,161	-0,783	0,994	-0,733	-0,989	-0,911	0,988	-0,917	-0,962	0,837	1	
Out..Activ.	-0,030	-0,730	0,947	-0,759	-0,950	-0,906	0,923	-0,931	-0,902	0,896	0,964	1

Figura 5. Diagrama de Scatterplots



A análise da matriz de correlações permitiu observar de que forma os setores de atividade se correlacionam e, assim, identificar grupos de setores cuja evolução afeta outros de forma proporcional direta ou inversa, conforme o sinal do coeficiente de correlação seja positivo ou negativo, respetivamente.

O setor da agricultura apresentou uma correlação moderada apenas com o setor das pescas num sentido positivo ($\rho=0,62$), o que permite equacionar que a variação do número de empresas nestes dois setores evoluiu no mesmo sentido de forma diretamente proporcional, em termos médios. A agricultura apresentou correlações muito baixas (inferiores a 0,3) com os restantes setores o que demonstra que a sua atividade está pouco relacionada com os outros setores.

O setor das pescas apresentou uma correlação fraca com o setor da eletricidade, gás e águas, e uma correlação forte e positiva com o setor da construção ($\rho=0,71$), do comércio, grosso e retalho ($\rho=0,84$), dos transportes e armazenagem ($\rho=0,75$) e do imobiliário ($\rho=0,65$), o que se interpretou como tendo origem no facto de estarem todos estes setores associados ao transporte de bens e mercadorias, tanto a nível interno como externo.

Contrariamente, o setor das pescas apresentou uma correlação negativa forte com a indústria transformadora ($\rho=-0,78$), assim como com os setores do alojamento e restauração ($\rho=-0,77$), da saúde e ação social ($\rho=-0,78$) e outras atividades ($\rho=-0,73$), enquanto que com a educação a sua correlação foi moderada negativa ($\rho=-0,57$).

À exceção dos casos particulares da agricultura e do setor das pescas, com diferentes sentidos da correlação, todos os restantes setores de atividade apresentaram correlações muito fortes, revelando em alguns casos uma relação inversa entre as variáveis e, noutros, uma relação direta. Assim sendo, não foi possível nesta fase distinguir um padrão de comportamento.

A partir da Figura 5 (diagrama de *scatterplots*) avaliou-se a linearidade das relações existentes entre os setores de atividade económica. O setor da indústria transformadora mantém uma relação linear positiva com os setores do alojamento e restauração, da educação, da saúde e ação social e o setor das outras atividades, o que significa que o aumento ou a diminuição da taxa de empresas no setor da indústria transformadora tende a estar associado a um aumento ou diminuição nos outros setores referidos. Isto é também visível pelos valores do Quadro 5

Contrariamente, o setor da indústria transformadora apresenta uma relação inversa com os setores de construção, de comércio a grosso e a retalho, transporte e armazenagem e da atividade imobiliária. O aumento do número de empresas do setor da indústria transformadora provoca uma

redução proporcional do número de empresas no setor da construção, assim como nos restantes setores referidos.

O setor da construção mantém também correlação linear positiva com os setores de comércio a grosso e a retalho, transporte e armazenagem e atividade imobiliária. Pode-se interpretar que o aumento do número de empresas na construção, motivada pela necessidade da reconstrução nacional, torna necessária a importação de matérias-primas, é acompanhado pelo aumento da atividade de comércio a grosso e a retalho para dar resposta à procura interna e, conseqüentemente, um aumento das empresas de transporte e armazenagem em termos médios. A atividade do setor de construção, do ponto de vista do número de empresas, implica também o desenvolvimento do setor do imobiliário uma vez que potencia a oferta no mercado imobiliário. Contrariamente a esta tendência, observou-se uma relação linear negativa entre o setor da construção e o setor indústria transformadora, o que pode ser explicado pelo facto do setor da indústria não estar ainda a beneficiar do desenvolvimento do setor da construção em termos da construção de infra-estruturas básicas de operacionalidade das indústrias transformadoras nacionais.

Observou-se uma relação linear positiva entre o setor do alojamento e a restauração e o setor imobiliário, bem como entre o primeiro e o do transporte e armazenagem. Neste caso pode-se interpretar que o setor do alojamento e restauração registou um desenvolvimento acentuado, sobretudo em 2010, em virtude da realização do Campeonato Africano de Futebol (CAN) e da preparação das infraestruturas (de alojamento e restauração) para a realização do Campeonato do Mundo de Hóquei em Angola que ocorreu em 2013. O setor do alojamento e a restauração apresentam uma relação positiva com o setor do transporte e armazenagem uma vez que utilizam os serviços externos permitindo um desenvolvimento dos transportes e logística.

Em relação à educação foi observada uma relação linear negativa com os setores do alojamento e restauração e da indústria transformadora, a qual pode eventualmente ser explicada pelo facto de Angola não ter neste momento uma oferta educativa acentuada orientada para estes setores face à procura. Com os restantes setores em análise (à exceção da agricultura), a educação apresentou uma relação linear positiva, o que revela que a formação de quadros nestas áreas beneficia o desenvolvimento de setores como a construção, o imobiliário, o comércio grosso e retalho, e o setor dos transportes.

Por sua vez, o setor da saúde e ação social revelou uma relação linear positiva com os setores da indústria transformadora e do alojamento e restauração, que pode ser associada ao facto destes dois últimos recorrerem a serviços de saúde e ação social. A relação linear positiva entre o setor da

educação e o setor da saúde e ação social pode ser compreendida pelo facto do primeiro permitir o desenvolvimento da saúde e ação social através da formação e capacitação de quadros.

1.4. Análise de Clusters

A contribuição mais expressiva para a aplicação do método de análise de *clusters* foi dada por Sokal e Sneath, no seu livro *Numerical Taxonomy. The principles and the practice of numerical classification* (1963) inicialmente aplicados à biologia. Dado um conjunto de n indivíduos, todos avaliados para p variáveis, tais indivíduos devem ser agrupados em *clusters* ou grupos, de forma que os mais semelhantes permaneçam no mesmo grupo. Geralmente, o número de classes não é conhecido *á priori*.

No caso em estudo, a análise de *clusters* pretende identificar as semelhanças e dissemelhanças entre os setores de atividade, de modo a formar grupos internamente homogéneos e externamente heterogéneos entre si. Para quantificar as semelhanças/dissemelhanças dos setores económicos, foram utilizadas as medidas de dissimilaridades baseadas na correlação, autocorrelação simples e parcial, no periodograma simples e integrado.

A dissimilaridade baseada na correlação calcula as distâncias entre os setores, com base no coeficiente de correlação de Pearson, introduzido anteriormente.

A dissimilaridade baseada na autocorrelação informa o quanto o valor de uma realização de uma variável aleatória é capaz de influenciar seus vizinhos (autocorrelação). A dissimilaridade baseada na autocorrelação simples é definida por,

$$DACF(X, Y) = \sqrt{(\hat{\rho}_X - \hat{\rho}_Y)^T \Omega (\hat{\rho}_X - \hat{\rho}_Y)},$$

em que $\hat{\rho}_i = (\hat{\rho}_{i,1}, \dots, \hat{\rho}_{i,m})$ é o vetor das estimativas dos coeficientes de autocorrelação da série temporal i , onde $\hat{\rho}_{i,k} = \frac{cov[X_t, X_{t+k}]}{\sigma^2}$, e Ω é uma matriz de pesos.

Quando $\Omega = I$ (matriz de identidade) a dissimilaridade baseada na autocorrelação, corresponde à distância euclidiana entre as matrizes de autocorrelação da série X e Y . Quando $\Omega = [cov(\hat{\rho})]^{-1}$ obtém-se a distância de Mahalanobis entre as autocorrelações.

No caso, o cálculo das distâncias pelas autocorrelações parciais é equivalente ao cálculo pela autocorrelação simples, no entanto, $\hat{\rho}_{i,k}$ são dados pela seguinte expressão:

$$\hat{\rho}_{i,k} = \frac{|P_k^*|}{|P_k|},$$

em que P_k^* é a matriz $k \times k$ de autocorrelações onde a última coluna da matriz P é substituída por $[\rho_1 \ \rho_2 \ \dots \ \rho_k]^T$. A matriz P é dada por:

$$\begin{bmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & 1 \end{bmatrix}$$

Para a dissimilaridade baseada no periodograma, as distâncias são calculadas com base no coeficiente do periodograma. Este coeficiente é utilizado para identificar a sazonalidade dos dados, ou seja, fenómenos que se repetem a cada período idêntico de tempo. A medida conhecida como logaritmo do periodograma normalizado (DNLP) é definida como a distância euclidiana entre os coeficientes dos periodogramas das séries X e Y ,

$$DLNP(X, Y) = \sqrt{\sum_{j=1}^{n/2} [P_X(w_j) - P_Y(w_j)]^2}$$

em que $P_X(w_j) = \frac{1}{n} |\sum_{t=1}^n x_t e^{-itw_j}|^2$ e $P_Y(w_j) = \frac{1}{n} |\sum_{t=1}^n y_t e^{-itw_j}|^2$ são os periodogramas das séries temporais X e Y , respetivamente, e $w_j = \frac{2\pi j}{n}$.

A medida de dissimilaridade entre duas séries temporais baseada na distância dos seus periodogramas integrados é definida por:

$$INT.PER(X, Y) = \int_{-\pi}^{\pi} |F_X(\lambda) - F_Y(\lambda)| d\lambda$$

em que :

F_X representa função de distribuição da série temporal X

F_Y representa a função de distribuição da série temporal Y

O conglomerado dos setores de atividade económica em estudo foi realizado através das técnicas de ligação hierárquicas que recorrem a passos sucessivos de agregação dos objetos considerados individualmente. Ou seja, cada objeto é um *cluster* e, esses *clusters* são sucessivamente agrupados de acordo com as suas proximidades (por isso estes métodos se dizem aglomerantes). Os métodos hierárquicos de conglomerado diferem no modo como as distâncias entre os objetos são calculadas (Marôco, 2011).

Um método de conglomerado é considerado hierárquico se ele consiste numa sequência de $(W + 1)$ conglomerados (G_0, G_1, \dots, G_W) , em que G_0 é a partição disjunta de todos os n indivíduos e G_W é a partição conjunta.

Na análise de *clusters* foram aplicados sete métodos hierárquicos de conglomerado, que a seguir se descrevem de forma sucinta.

No método de ligação simples (*single linkage*), ou critério do vizinho mais próximo, a distância entre dois grupos (N e M) é determinada pela menor distância (d_{ij}) entre os objetos mais parecidos entre os dois grupos. Isto é:

$$D_{NM} = \min_{\substack{i \in N \\ j \in M}} \{d_{ij}\}$$

Pelo método de ligação completa (*complete linkage*), ou do vizinho mais distante, a distância entre dois grupos é definida como a sendo a dos dois vizinhos menos parecidos, ou seja, formam-se todos os pares com um membro de cada grupo, e a distância entre os grupos é definida pelo par que se encontra mais distante. Este método, geralmente, resulta na formação de grupos compactos e discretos, sendo os seus valores de similaridade relativamente pequenos (Bussab, Miazaki & Andrade, 1990). A distância entre os dois grupos é definida por:

$$D_{NM} = \max_{\substack{i \in N \\ j \in M}} \{d_{ij}\}$$

No método de ligação média (*average linkage*) é usada a similaridade média dos indivíduos ou grupo que se pretende unir a um grupo já existente. Nos métodos de conglomerado com base na média aritmética, é calculada a média de todas as distâncias entre todos os pares de elemento dos dois grupos. A distância entre os dois grupos é definida por:

$$D_{NM} = \frac{1}{|N| |M|} \sum_{\substack{i \in N \\ j \in M}} d_{ij}$$

O método de ligação de centróide (*centroid*) considera a distância entre os centróides, ou seja, a distância entre os dois grupos é calculada com base na distância entre a média de todas as variáveis desses dois grupos (Marôco, 2011). A distância entre os dois grupos é definida por:

$$D_{NM}^c = d(\bar{X}_N, \bar{Y}_M)$$

onde $\bar{X}_N = \frac{1}{|N|} \sum_{i \in N} x_i$ e $\bar{Y}_M = \frac{1}{|M|} \sum_{j \in M} y_j$ são os centróides dos *clusters* N e M.

O método de Ward (*ward linkage*) considera a distância entre dois grupos através do método de ligação de centróide sendo definida por:

$$D_{NM} = \frac{|N| |M| D_{NM}^{c^2}}{|N| + |M|}$$

onde D_{NM}^c é a distância entre N e M pelo método de centróide.

Com o método de ligação do McQuitty (*McQuitty linkage*), quando dois conglomerados (K, L) são unidos, a distância entre o novo conglomerado (M) e qualquer outro conglomerado é calculada com base na média das distâncias dos conglomerados a serem unidos para esse outro conglomerado (McQuitty, 1964). A distância é definida por:

$$D_{(KL)M} = \frac{D_{KL} + D_{LM}}{2}$$

No método de ligação da mediana (*median linkage*) entre dois clusters, é a mediana das distâncias entre todos os possíveis pares de observações (i, j) com $i \in N$ e $j \in M$. A distância entre os dois grupos é definida por :

$$D_{NM} = \text{mediana} \{d_{ij} : i \in N, j \in M\}$$

Depois de obtidos os conglomerados através dos diferentes métodos de ligação hierárquica, recorreu-se ao coeficiente de correlação cofenética para avaliar a qualidade de ajustamento entre as dissimilaridades e os resultados gerados pelos métodos de conglomerado (Sokal & Rohlf, 1962).

A correlação cofenética mede o grau de ajuste entre a matriz de similaridade original (matriz S) e a matriz resultante da simplificação proporcionada pelo método de conglomerado (matriz C). A matriz C é a obtida após a construção do dendrograma. Tal correlação foi calculada com a seguinte equação:

$$r_{cof} = \frac{\sum_{i=1}^{p-1} \sum_{j=i+1}^p (c_{ij} - \bar{c})(s_{ij} - \bar{s})}{\sqrt{\sum_{i=1}^{p-1} \sum_{j=i+1}^p (c_{ij} - \bar{c})^2} \sqrt{\sum_{i=1}^{p-1} \sum_{j=i+1}^p (s_{ij} - \bar{s})^2}}$$

em que: c_{ij} : valor de similaridade entre os indivíduos i e j , obtidos a partir da matriz cofenética;

s_{ij} : valor de similaridade entre os indivíduos i e j , obtidos a partir da matriz de similaridade;

$$\bar{c} = \frac{2}{p(p-1)} \sum_{i=1}^{p-1} \sum_{j=i+1}^p c_{ij}$$

$$\bar{s} = \frac{2}{p(p-1)} \sum_{i=1}^{p-1} \sum_{j=i+1}^p s_{ij}$$

O coeficiente de correlação cofenética equivale à correlação de Pearson entre a matriz de similaridade original e aquela que é obtida após a construção do dendrograma. Ou seja, é a medida que indica o quanto o dendrograma (que é um modelo do comportamento de similaridade) modela o comportamento real. Assim, quanto mais próximo de 1, menor será a distorção provocada pelo conglomerado dos indivíduos através deste método (Bussab et al., 1990). A matriz cofenética é a

matriz das distâncias, na qual as distâncias entre os pares originais de objetos são substituídas pelas distâncias calculadas entre os grupos, no momento da formação dos conglomerados (Carr, Young, Aster & Zhang, 1999).

O Quadro 6 apresenta os valores do coeficiente de correlação cofenética para os critérios de agregação que apresentaram os dois melhores resultados e, para cada critério de agregação, as medidas de dissimilaridade baseadas na correlação, no periodograma simples, na autocorrelação parcial, autocorrelação simples e no periodograma integrado, bem como o critério de agregação que pior resultado gerou para as medidas de dissimilaridades analisadas.

Quadro 6. Coeficiente de correlação cofenética

Crítérios de Agregação	Medida de Dissimilaridade	Coeficiente de correlação cofenético
<i>Complete Linkage</i>	Periodograma Simples	0,95316
	Correlação	0,94902
	Autocorelação Parcial	0,85102
	Autocorelação Simples	0,69785
	Periodograma Integrado	0,67323
<i>Average Linkage</i>	Correlação	0,97331
	Periodograma Simples	0,95759
	Autocorelação Parcial	0,89994
	Autocorelação Simples	0,89503
	Periodograma Integrado	0,67947
<i>Single Linkage</i>	Correlação	0,97166
	Periodograma Simples	0,93073
	Autocorelação Simples	0,88026
	Autocorelação Parcial	0,86681
	Periodograma Integrado	0,48853
<i>Ward Linkage</i>	Correlação	0,92995
	Autocorelação Parcial	0,85754
	Periodograma Simples	0,81315
	Autocorelação Simples	0,69181
	Periodograma Integrado	0,65003

Com base no Quadro 6 observa-se que foi com o critério de agregação *average linkage* que se registraram maiores coeficientes de correlação cofenética em todas as medidas de dissimilaridade,

tendo sido a medida baseada na correlação a que registou o melhor resultado (0,97331) seguindo-se as dissimilaridades obtidas pelo periodograma simples (0,95759), também as dissimilaridades obtidas com base na autocorrelação parcial e na autocorrelação simples apresentaram valores altos do coeficiente cofenético de (0,89994 e 0,89503 respetivamente). O periodograma integrado foi o que registou um menor coeficiente cofenético (0,67947).

Os segundos melhores resultados das medidas de dissimilaridades registaram-se em diferentes criterios de agregação. As dissimilaridades obtidas pela correlação obtiveram o segundo melhor resultado do coeficiente de correlação cofenética através do critério *single linkage* (0,971166). O periodograma simples registou o segundo melhor coeficiente no critério *complete linkage* (=0,95316). A autocorrelação parcial registou o segundo melhor coeficiente cofenético com o critério de agregação *centroid linkage* (0,88813), enquanto que para a autocorrelação simples o segundo melhor coeficiente foi obtido através do critério de agregação *median linkage* (0,89356). Por fim, o periodograma integrado apresentou o segundo melhor valor do coeficiente de correlação cofenética através do critério *McQuitty linkage* (0,67847).

O critério de agregação *ward linkage* foi o que gerou piores resultados, quer para as dissimilaridades obtidas pela correlação (0,92995) quer para as dissimilaridades obtidas pelo periodograma simples (0,81315).

Para prosseguir com a análise de *clusters*, recorreu-se ao critérios de agregação através da média e foram consideradas as duas medidas de dissimilaridades com melhores resultados do coeficiente de correlação cofenético, a correlação e o periodograma simples. No sentido de proceder a uma análise comparativa, analisaram-se as dissimilaridades obtidas através do critério de agregação *average linkage* e através do critério de agregação com o segundo melhor resultado para a correlação e o periodograma simples, que foram o critério *single linkage* e *complete linkage* respetivamente.

Os dendrogramas obtidos além de permitirem obter uma representação gráfica das distâncias entre as diferentes observações permitem também observar em que nível de similaridade os grupos são formados, assim como a composição dos conglomerados da partição final. Além disso, permite também ter uma visualização, ao longo do processo de conglomeração, de quais os grupos que se vão formando.

Figura 6. Dissimilaridade baseada na correlação

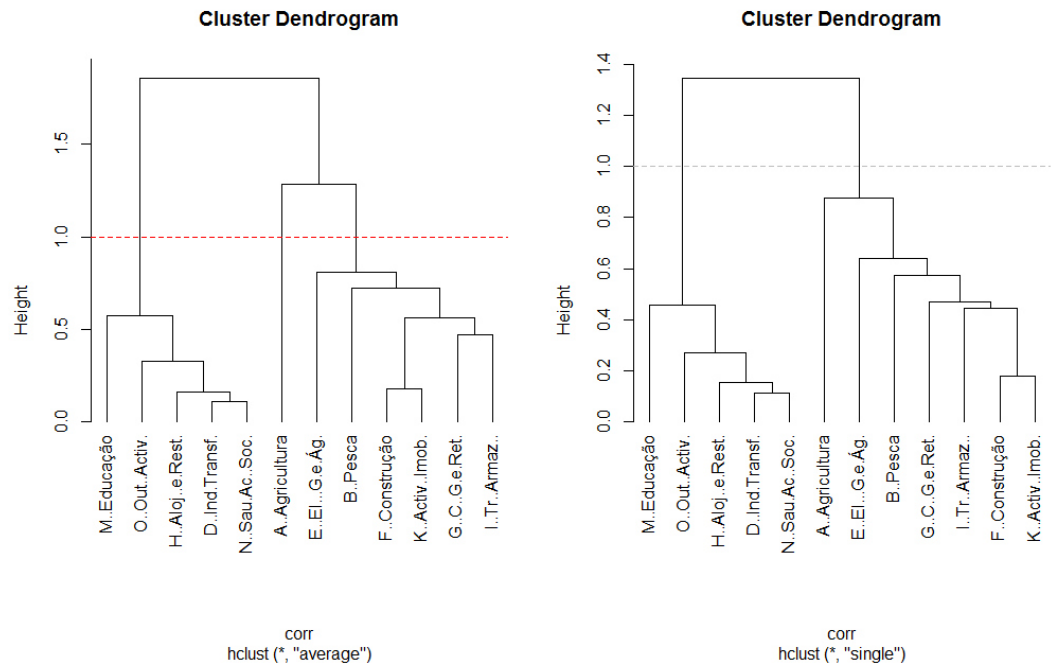
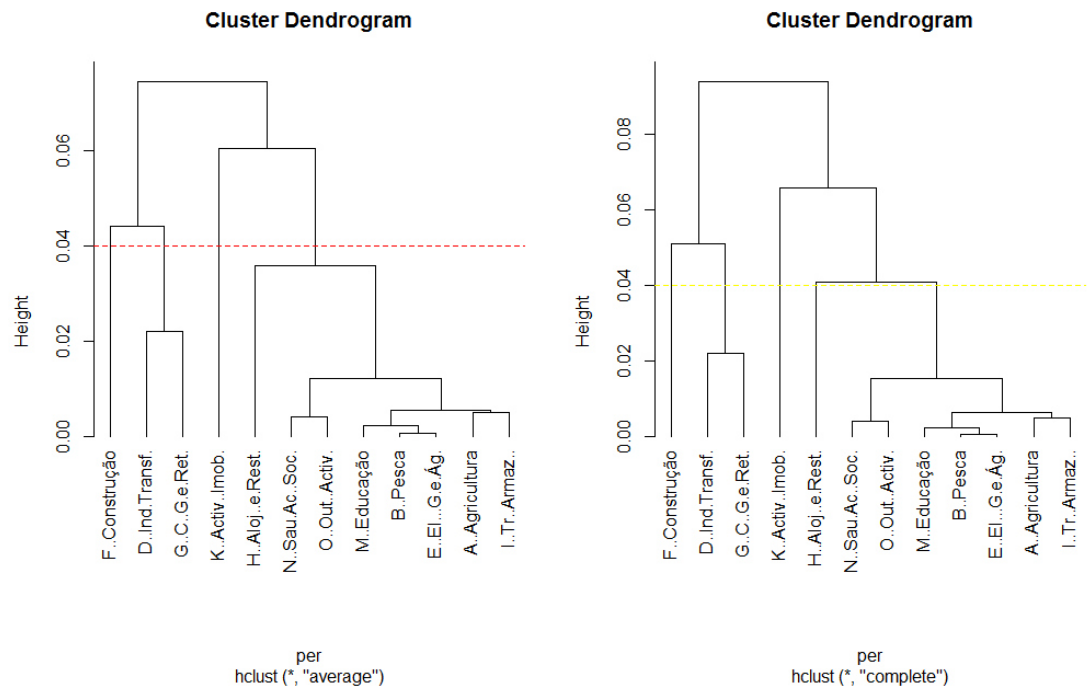


Figura 7. Dissimilaridade baseada no periodograma simples



O Quadro 7 permitiu sintetizar a informação obtida pelos métodos representados no dendrograma da Figura 6 e da Figura 7.

Quadro 7. Sumário dos clusters/grupos obtidos

Dendrograma		Nº de clusters	Nº de setores em cada cluster	Grupos/Clusters
Figura 6. Dissimilaridade baseada na correlação	Critério <i>Average Linkage</i>	3	A ₁ : 5	Educação, Out.Ativ. Aloj.e Rest., Ind. Transf., Saud. Ac.Soc.
			A ₂ : 1	Agricultura
			A ₃ : 6	El.G. e Ág., Pesca, Construção, C.G.e Ret., Activ. Imob, Tr. Armaz.
	Critério <i>Single Linkage</i>	2	B ₁ : 5	Educação, Out.Ativ., Aloj.e Rest., Ind.Transf., Saud.Ac.Soc.
			B ₂ : 7	Agricultura, El.G. e Ág., Pesca, C.G.e Ret., Tr.Armaz., Construção, Activ.Imob.
Figura 7. Dissimilaridade baseada no periodograma simples	Critério <i>Average Linkage</i>	4	C ₁ : 1	Construção
			C ₂ : 2	Ind.Transf., C.G.e Ret.
			C ₃ : 1	Activ.Imob.
			C ₄ : 8	Aloj. e Rest, Saud.Ac.Soc, Out.Ativ. Educação, El.G e Ag., Pesca, Agricultura, Tr.Armaz.
	Critério <i>Complete Linkage</i>	5	D ₁ : 1	Construção
			D ₂ : 2	Ind.Transf., C.G.e Ret.
			D ₃ : 1	Activ.Imob.
			D ₄ : 1	Aloj. e Rest.
			D ₅ : 7	Saud.Ac.Soc., Out.Ativ., Educação, Pesca, El.G e Ag., Agricultura, Tr.Armaz.

Na Figura 6 o corte do dendrograma a uma altura de 1 revelou, para o caso em que se usou o critério *average linkage* e a medida de correlação, a formação de três grupos de setor de atividade económica. O primeiro grupo (A₁) integrou os setores da educação, outras atividades, alojamento e restauração, indústria transformadora, e saúde e ação social. O segundo grupo (A₂) foi formado apenas pelo setor agrícola, o que sugere a ausência de semelhanças com os outros setores. O terceiro conglomerado (A₃) foi constituído por seis setores económicos (a eletricidade, gás e água, as pescas, a construção, a atividade imobiliária, o comércio grosso e retalho e os transportes e armazenagem).

Os resultados de conglomerado pelo critério *average linkage* e dissimilaridades baseada na correlação foram ao encontro dos resultados observados através da matriz de correlações. O setor da agricultura apresentou correlações quase nulas com os restantes setores de atividade, pelo que forma um grupo isolado da análise de conglomerados. Os setores que integraram o terceiro *cluster*, com exceção da relação entre as pescas e a eletricidade, gás e água, apresentaram correlações positivas e fortes (superiores a 0,7) entre si, o que nos leva a supor a existência de uma evolução proporcional e de igual sentido do conjunto dos setores quando um deles é estimulado ou afetado.

O critério *single linkage*, através da medida de dissimilaridades baseada na correlação, permitiu identificar dois grupos. O primeiro grupo obtido (B_1) integrou os setores da educação, outras atividades, alojamento e restauração, indústria transformadora, saúde e ação social. O setor agrícola, que no método anterior ficou num grupo isolado (A_2), surgiu nesta análise alocado juntamente com os setores da eletricidade gás e água, pesca, comércio a grosso e a retalho, transporte e armazenamento, construção e atividade imobiliária no grupo (B_2).

No que se refere às dissimilaridades com base na correlação, o critério *single linkage* permitiu obter conglomerados de constituição muito semelhante aos obtidos pelo critério *average linkage*. O primeiro grupo obtido através do critério *single linkage* (B_1) coincide com o primeiro grupo do critério *average linkage* (A_1), enquanto o segundo *cluster* do critério *single linkage* (B_2) integrou os setores do primeiro e do segundo grupo obtidos pelo critério *average linkage* (A_1 e A_2). A formação dos *clusters* foi assim influenciada pelas correlações observadas entre os setores de atividade, sendo que os que apresentaram correlações fortes e positivas integraram o mesmo grupo, diferente do grupo dos setores cuja relação se observou inversa.

Através da Figura 7, observa-se que o corte do dendrograma a uma altura de 0,04 permitiu a constituição de diferentes grupos quando utilizado o critério *average linkage* e o critério *complete linkage*.

No caso do critério *average linkage*, o quarto grupo (C_4) é parcialmente semelhante ao grupo B_2 obtido pelo critério *single linkage* e medida de dissimilaridade com base na correlação. O setor de construção e da atividade imobiliária formam dois grupos (C_1 e C_3), enquanto o setor da indústria transformadora e do comércio a grosso e retalho formam em conjunto um outro grupo (C_2). O quarto *cluster* (C_4) obtido pelo critério *average linkage* e a medida do periodograma simples foi o que integrou

o maior número de setores, no total de oito (alojamento e restauração, saúde e ação social, outras atividades, educação, pescas, eletricidade, gás e água, agricultura, transportes e armazenagem). O critério *complete linkage* permitiu identificar cinco grupos. O grupo C₄ obtido através do critério *average linkage* é subdividido em dois grupos com o critério *complete linkage*, constituindo o grupo D₄ (que integra apenas o alojamento e restauração) e os outros setores ficam agrupados no D₅. Os restantes três grupos obtidos foram iguais aos que resultaram do critério *average linkage*.

De salientar que para os dendrogramas baseados no periodograma, a análise dos grupos não pode ser comparada com a matriz de correlações, porque este método não contabiliza as relações lineares entre os setores, considera apenas possíveis as semelhanças dos comportamentos sazonais.

O segundo (A₂) e o terceiro (A₃) conglomerados obtidos através do critério *average linkage* e da medida de correlação permitiram confirmar algumas das suposições equacionadas na análise descritiva. O segundo grupo (A₂) (à excepção do setor das pescas e da eletricidade, gás e água) confirmou a suposição de que os setores da construção, comércio, grosso e retalho, transportes e armazenagem e a atividade imobiliária integrariam um mesmo *cluster*, uma vez que apresentaram uma tendência crescente da taxa de distribuição do número de empresas. O terceiro conglomerado (A₃) integrou no mesmo *cluster* os setores que apresentaram uma tendência decrescente do número de empresas, designadamente, os setores da indústria transformadora, alojamento e restauração, saúde e ação social e outras atividades; o mesmo conglomerado integrou ainda o setor da educação que na análise descritiva não revelou uma tendência semelhante aos demais setores.

A análise da taxa de distribuição do número de empresas por setor de atividade económica revelou também um comportamento semelhante entre os setores da agricultura, das pescas e da eletricidade, gás e água, resultado que apenas foi confirmado através do segundo *cluster* do critério *single linkage* (B₂) e pelo quarto *cluster* do critério *average linkage* (C₄).

II. A Experiência do Banco de Portugal na Produção de Estatísticas sobre Sociedades Não Financeiras

A informação estatística das Sociedades Não Financeiras (SNF) em Portugal tem sido publicada desde 1983, altura em que as empresas eram inquiridas a partir do Inquérito Anual da Central de Balanços (IACB) que não abrangia todos os setores de atividade das SNF. O IACB cobria cerca de 17 mil empresas e a recolha da informação era realizada de forma quase automática. Nos dias de hoje é utilizado o Inquérito Trimestral às Empresas Não Financeiras (ITENF) e a Informação Empresarial Simplificada (IES) que adiante se explicitam.

2. Enquadramento Legal

O Banco de Portugal (BdP) publica as estatísticas das SNF ao abrigo da Lei Orgânica do Banco de Portugal (Lei nº 5/98, de 31 de janeiro) e suas alterações subsequentes, do Decreto-Lei nº 8/2007, de 17 de janeiro, que cria a IES, da Lei do Sistema Estatístico Nacional (Lei nº 22/2008 de 13 de maio), da Lei Orgânica do Instituto Nacional de Estatística (Decreto-Lei n.º 136/2012, de 2 de julho), do Decreto-Lei nº 381/2007 (de 14 de novembro) que substituiu a CAE-Rev.2.1, a partir de 1 de janeiro de 2008, pela Classificação Portuguesa de Atividades Económicas, Revisão 3 (CAE-Rev.3), e a Recomendação da Comissão Europeia, de 6 de maio de 2003 (2003/361/CE) relativa à definição de micro, pequenas e médias empresas.

De acordo com o artigo 13.º da Lei Orgânica do Banco de Portugal, *“competete ao Banco a recolha e elaboração das estatísticas monetárias, financeiras, cambiais e da balança de pagamentos, designadamente no âmbito da sua colaboração com o Banco Central Europeu (BCE), o Banco pode exigir a qualquer entidade, pública ou privada, que lhe sejam fornecidas diretamente as informações necessárias para cumprimento do estabelecido no número anterior ou por motivos relacionados com as suas atribuições.”* (BdP, 2015).

A produção das estatísticas das SNF constitui um importante instrumento de apoio à tomada de decisão, na medida em que permite a análise da evolução económica e financeira dos setores de atividades económica e, por esta via, orientar a adoção e implementação de medidas e políticas públicas necessárias ao desenvolvimento dos diferentes setores.

3. Fontes de Informação e Controlo de Qualidade

Para compilação das estatísticas das SNF, o BdP utiliza duas fontes de informação o Inquérito Trimestral às Empresas Não Financeiras e a Informação Empresarial Simplificada. Para melhor compreender o que é o ITENF e a IES apresenta-se de seguida a descrição de cada um (BdP, 2013).

3.1. Inquérito Trimestral às Empresas não Financeiras

O ITENF é uma operação estatística da responsabilidade conjunta do Banco de Portugal e do Instituto Nacional de Estatística (INE) de Portugal, com base num protocolo interinstitucional que minimiza a carga estatística sobre as empresas inquiridas. Neste contexto, e ao abrigo deste protocolo, o (INE) de Portugal recolhe a informação e mantém o contacto direto com as empresas para esclarecimento de eventuais dúvidas sobre os dados micro.

O ITENF tem sido realizado desde 2000 e tem como principal objetivo a recolha de informação trimestral junto de cerca de 4 mil empresas selecionadas com base numa amostra aleatória estratificada. As empresas selecionadas para serem inquiridas trimestralmente estão sujeitas a critérios predefinidos da atividade económica tais como o volume de negócios, o total de ativos, as exportações e as importações. A seleção das empresas é levada a cabo pelo INE de Portugal e a informação recolhida incide, essencialmente, sobre um conjunto de variáveis de natureza contabilística relativas à atividade e à situação financeira das empresas (BdP, 2013).

O processo de validação da informação de base é semelhante ao da IES, diferenciando-se apenas no contacto com os reportantes, sendo as questões da central de balanços remetidas ao INE para esclarecimento junto das empresas.

3.2. Informação Empresarial Simplificada

A IES foi formalmente criada pelo Decreto-Lei N° 8/2007, de 17 de janeiro, e incide sobre o universo das SNF. A IES é de resposta obrigatória, tendo o seu registo passado para o formato automático, o que impulsionou a redução do prazo de submissão da informação de doze meses para seis meses e meio. Este documento veio, assim, alargar o âmbito da inquirição da informação a todos os setores de atividade, garantindo a representatividade das atividades desenvolvidas pelo universo

das SNF. Atualmente, os dados recolhidos pela IES reportam-se ao exercício económico de um ano, num universo de 370 mil empresas. A IES permite que, através de uma única submissão eletrónica, as empresas cumpram as suas obrigações de prestação de contas anuais junto de quatro entidades públicas: Autoridade Tributária, Ministério da Justiça, INE e BdP (BdP, 2014).

A informação anual é reportada ao abrigo do Decreto-Lei n.º 8/2007, de 17 de janeiro, e agrega, num único ato, o cumprimento das seguintes obrigações: entrega da declaração anual de informação contabilística e fiscal ao Ministério das Finanças; registo da prestação de contas junto das conservatórias do registo comercial (n.º 1 do artigo 15.º do Código do Registo Comercial); prestação de informação de natureza estatística ao INE (n.º 1 do artigo 6º da Lei do Sistema Estatístico Nacional); prestação de informação relativa a dados contabilísticos anuais para fins estatísticos ao BdP (artigo 13.º da Lei Orgânica do BdP).

A informação recolhida através da IES é sobretudo de natureza contabilística. A IES permite a recolha de variáveis que incluem elementos de caracterização geral da empresa, informação contabilística contida nas principais demonstrações financeiras previstas no Sistema de Normalização Contabilística (SNC) e ainda um conjunto adicional de informação sobre a atividade e a situação das empresas, como o número de pessoas ao serviço e as transações de bens e de serviços com o exterior, recolhidas para efeitos estatísticos.

À semelhança da validação de outras informações de base, a IES valida a coerência da informação contabilística que recolhe, através de uma validação temporal (análise da evolução temporal, variações em cadeia e homólogas, e controlo de *outliers*), de uma validação interna (coerência dos dados entre si) e de uma validação externa (confronto com outras bases de dados ao nível da informação base).

3.3. Controlo de Qualidade

O controlo da qualidade e a análise da consistência da informação estatística desempenham um papel fulcral e revestem-se de extrema importância na obtenção de estatísticas com qualidade e fiabilidade, uma vez que o processo de validação da informação proveniente das SNF reduz substancialmente as inconsistências dos dados reportados.

O controlo de qualidade da informação incorpora o controlo automático, no qual são validadas as regras de submissão das declarações da IES e do ITENF, a validação com as fontes internas e externas e por último, a análise técnica.

Para efeitos de validação da informação, a Central de Balanços utiliza internamente a informação da Central de Responsabilidades de Crédito (CRC), do Sistema Integrado de Estatísticas de Títulos (SIET), da Balança de Pagamentos (BoP) e da Posição de Investimento Internacional (PII), a Tabela de Entidades, designada por Sistema de Partilha de Informação de Referência (SPAI) e como fontes externas a Segurança Social (SS) e a Autoridade Tributária e Aduaneira (ATA).

A natureza da informação básica diz respeito a micro dados que são organizados numa base de dados permitindo, pelo seu nível de detalhe, um controlo mais eficaz de qualidade estatística, não só a nível do conteúdo da informação, como também a nível dos processos de agregação a que uma informação está sujeita.

Os critérios de seleção das empresas para análise têm em consideração a utilização da informação para fins estatísticos e a análise da situação económica e financeira das sociedades não financeiras. No que diz respeito ao ITENF, o INE realiza previamente um procedimento de controlo de qualidade que consiste, por um lado, na identificação de *outliers* através da aplicação do método *Hidiroglou-Berthelot* (1986, cit. in Belcher, 2003) a determinadas variáveis submetidas e por outro, no contacto com a empresa, para confirmação e/ou correção de eventuais evoluções anómalas e/ou erros no reporte.

4. Compilação dos Dados Extrapolados

O procedimento de extrapolação dos dados é utilizado para obter estimativas para todas as SNF das variáveis de interesse, com base na amostra do ITENF.

4.1. Tratamento da Informação Anual

O tratamento da informação anual reveste-se de extrema importância para os processos de extrapolação e de conciliação entre dados anuais e trimestrais. Para simplificar o processo e garantir uma maior qualidade da informação, os inquéritos dirigidos às empresas assentam em informação que as mesmas produzem, evitando assim, a realização de cálculos adicionais para o preenchimento dos inquéritos. Com a introdução da IES em 2007, foi solicitado às empresas a informação contabilística

dos anos de 2005 e 2006, para permitir a comparabilidade entre os dados. Contudo, o incumprimento por parte de algumas SNF da obrigatoriedade de reporte dos dados e o não cumprimento dos prazos estabelecidos, leva a existência de lacunas nos dados, isto é, dados omissos (não resposta). Quando são detetadas empresas com dados omissos é feita uma estimativa para esses dados a partir da informação disponível no universo de referência do setor das SNF e nas bases de dados geridas pelo BdP (fontes internas como a CRC e o SIET).

O tratamento de dados omissos tem como finalidade a obtenção de estimativas dos valores das variáveis do balanço e da demonstração de resultados das empresas que não submeteram a informação à IES.

Para efeitos de tratamento da informação anual recorre-se a diversos métodos de imputação da não resposta (dados omissos). Neste contexto, será oportuno descrever os métodos de tratamento da não resposta utilizados pelo BdP, nomeadamente o método Cold-Deck (Little, R. J. A. 1988) e o método de imputação pela média.

O método Cold-Deck é um método de imputação que preenche o valor em falta a partir do valor observado para a própria empresa na base dados da CB, ou numa fonte alternativa (por exemplo numa outra base de dados ou na mesma mas num período diferente). O valor observado na fonte alternativa pode ser diretamente imputado ao valor em falta, ou pode ser-lhe aplicada uma tendência de crescimento, caso seja relativo a um período diferente.

A imputação do valor em falta para a empresa é feita com base em dados de um período anterior ou posterior ao período em análise, sobre os quais se aplica a taxa média de variação (total de ativo ou volume de negócios) das empresas com resposta no mesmo estrato da classificação da atividade económica e volume de negócios.

A imputação é efetuada de forma sequencial por forma a resolver-se todas as situações de dados omissos.

O tratamento de não resposta segundo a imputação pela média comporta quatro etapas aplicadas de forma sequencial. As três primeiras aplicam-se de forma sucessiva às empresas para as quais a etapa anterior não permitiu obter estimativas. A última é aplicada a todas as empresas e trata da calibragem da estimativa.

Na 1ª etapa de imputação pela média, são imputados a cada empresa sem resposta os valores médios observados nas empresas do mesmo estrato de CAE, volume de negócio, total do ativo E QUE cumpram os requisitos predefinidos (BdP, 2013). Na 2ª etapa, são imputados (para as empresas as quais não foram estimados valores na etapa 1) os valores médios observados com base na informação das empresas do mesmo estrato de CAE, volume de negócios e total do ativo. Na 3ª etapa é repetido o mesmo procedimento. Por último, na etapa de calibragem os valores imputados a cada empresa, em cada uma das etapas anteriores, são multiplicados por um ponderador, obtido de forma a garantir a estabilidade da classificação original da empresa, no que concerne ao volume de negócios e total do ativo (BdP, 2013). A estimativa final obtém-se usando uma das seguintes fórmulas:

$$Pond_i = \frac{vn_i}{vn_{i^*}} \quad \text{ou} \quad Pond_i = \frac{ta_i}{ta_{i^*}}$$

A primeira fórmula é usada na demonstração de resultados, enquanto que a segunda se usa no balanço.

$$y_i^f = y_i^* \times Pond_i$$

onde:

- vn_{i^*} - Valor imputado para o volume de negócios da empresa i ;
- vn_i - Volume de negócios da empresa i no universo de referência;
- ta_{i^*} - Valor imputado para o total do ativo da empresa i ;
- ta_i - Total do ativo da empresa i no universo de referência;
- $Pond_i$ - Ponderador da empresa i ;
- y_i^* - Valor imputado para a variável y para a empresa i ;
- y_i^f - Estimativa final da variável y para a empresa i .

4.2. Extrapolação de Dados Trimestrais

O procedimento de extrapolação dos dados trimestrais é antecedido pelo seu tratamento, processo que permite constituir a amostra do ITENF para variáveis de natureza contabilística, mediante a utilização de distintos critérios de estratificação: o volume de negócio, o total de ativos, as exportações e importações, e as Sociedades Gestoras de Participações Sociais (SGPS), (BdP, 2013).

Em Portugal, como os resultados totais para as empresas da base de amostragem não diferem substancialmente, comparativamente ao universo das SNF, no que concerne ao total do ativo e ao volume de negócios, os dados obtidos na amostragem são utilizados para a extrapolação do universo. Para calibrar o total amostral da variável de interesse e melhorar a precisão das estimativas, utiliza-se a informação anual a partir do estimador pelo rácio, com vista a reduzir ou eliminar um potencial enviesamento da estimativa. O estimador pelo rácio é um parâmetro estatístico e é definido como sendo a relação da média de duas variáveis (BdP, 2013) sendo dado pela seguinte expressão:

$$\hat{\tau}_{g^k} = \frac{\hat{y}_{g^k}}{\hat{x}_{g^k}} X_{g^k}$$

onde:

\hat{y}_{g^k} é o total amostral da variável de interesse y para o pós-estrato g^k , dado por $\sum_{i=1}^{n_{g^k}} \frac{y_{g^k i}}{\pi_i}$

\hat{x}_{g^k} é o total amostral da variável auxiliar x para o pós-estrato g^k , dado por $\sum_{i=1}^{n_{g^k}} \frac{x_{g^k i}}{\pi_i}$

X_{g^k} é o total populacional conhecido da variável auxiliar, dado por $\sum_{j=1}^{N_{g^k}} x_{g^k j}$

π_i é a probabilidade de seleção da empresa i

g^k pós-estrato – estrato a que a empresa pertence com base no setor de atividade e dimensão

A estimativa do total populacional da variável de interesse corresponde à soma das estimativas obtidas para os pós-estratos:

$$\tau = \sum_{g^k=1}^{G^k} \hat{\tau}_{g^k}$$

Utilizam-se como variáveis auxiliares o total de rendimentos, o total de ativo, as exportações e as importações (BdP, 2013).

4.3. Conciliação dos Dados Anuais e Trimestrais

A utilização de uma fonte de informação de periodicidade anual e outra trimestral provou existirem algumas diferenças nos valores obtidos, pelo que se recorreu ao procedimento de conciliação designado por *benchmarking* para anular as mesmas.

De uma forma geral, os modelos de *benchmarking* garantem o alisamento das séries finais, a consistência temporal entre as séries trimestrais ajustadas e as séries anuais, e o equilíbrio contabilístico (Denton, 1971).

As séries trimestrais finais são as séries que resultam do *benchmarking* e que representam as estimativas trimestrais finais da Central de Balanços. Estas séries apresentam algumas propriedades:

- As séries trimestrais são consistentes com o valor anual (isto é, são 'séries ajustadas' ao valor anual);
- Observam o equilíbrio contabilístico em todos os trimestres;
- O *benchmarking* atenua a variabilidade das séries extrapoladas, o que permite corrigir os desníveis artificiais entre valores de amostras diferentes.

O modelo utilizado tem como objetivo a minimização do somatório das variações entre as séries trimestrais finais (não observadas) e as séries estruturais, sujeitas por um lado às restrições de igualdade entre os dados anuais e trimestrais finais e por outro às restrições contabilísticas.

$$\min_y \left[\sum_{s=1}^S \sum_{t=1}^T \left(\frac{y_{s,t} - T_{s,t}}{T_{s,t}} \right)^2 \right] \quad (1)$$

onde:

$T_{s,t}$ - Valor trimestral estrutural da variável s no trimestre t

$y_{s,t}$ - Valor trimestral final da variável s no trimestre t

s.a:

$$\begin{cases} y_{t,s} = y_{0,a,s} \text{ para } t = 4 \times a \text{ e } s = 1, \dots, S_1, \text{ variável de stock} \\ \sum_{t=4 \times (a-1)+1}^{4 \times a} y_{t,s} = y_{0,a,s} \text{ para } s = S_1 + 1, \dots, S, \text{ variável de fluxo} \\ \sum_{s=1}^{S_1} (q_s \times y_{t,s}) + \sum_{s=S_1+1}^S \left(q_s \times \sum_{t=4 \times (a-1)+1}^t y_{t,s} \right) = 0 \end{cases} \quad (2)$$

onde:

$y_{t,s}$ - Variável trimestral final

$y_{0\ as}$ - Valor anual da variável s no ano a

q_s - Coeficiente da variável s na expressão de equilíbrio contabilístico

a - Número de observações das séries anuais;

T - Número de observações das séries trimestrais ($T \geq 4 \times a$)

S - Número de variáveis estimadas pelo modelo = S_1 Variáveis de stock + $(S - S_1)$ variáveis de fluxo

Neste problema, é importante interpretar as restrições descritas (2), pois as duas primeiras dizem respeito às restrições temporais que são definidas para variáveis de stock ou fluxo.

As variáveis de *stock* correspondem às variáveis do balanço, nomeadamente as posições em final de período, como por exemplo os financiamentos obtidos. No que concerne a estas variáveis, o quarto trimestre deve ser igual ao valor anual, dado pela seguinte expressão:

$$y_{t,s} = y_{0\ as} \text{ para } t = 4 \times a$$

As variáveis de fluxo correspondem à demonstração de resultados, como por exemplo o volume de negócios. Assim, o fluxo acumulado dos quatro trimestres deve ser igual ao valor anual e é dado por:

$$\sum_{t=4 \times (a-1)+1}^{4 \times a} y_{t,s} = y_{0\ as} \text{ para } s = S_1 + 1$$

A terceira restrição do problema é a restrição contabilística, que é aplicada apenas aos valores trimestrais e impõe que os valores ajustados mantenham o equilíbrio contabilístico em cada trimestre, sendo dada por:

$$\sum_{s=1}^{S_1} (q_s \times y_{t,s}) + \sum_{s=S_1+1}^S \left(q_s \times \sum_{t=4 \times (a-1)+1}^t y_{t,s} \right) = 0$$

Como resultado da conciliação, das 3 restrições obtém-se a minimização dos desníveis artificiais chegando à coincidência dos dados trimestrais e anuais. Estes dois procedimentos (a conciliação dos dados anuais e trimestrais e a extrapolação dos dados trimestrais) são abordados com detalhe no suplemento ao Boletim Estatístico do BdP (BdP, 2013).

5. Produtos e Serviços Estatísticos da Central de Balanços do Banco de Portugal

A Central de Balanços do BdP disponibiliza produtos estatísticos como os Quadros do Setor (QS), os Quadros da Empresa e do Setor (QES), a Série Longa dos Quadros do Setor (SLQS), *Bank for Accounts of Companies Harmonised* (BACH) e as Estatísticas do Boletim Estatístico. De salientar, que a produção destes serviços, permite ainda a análise económica, investigação, obtenção de indicadores para avaliar a estabilidade financeira, obtenção de informação para suporte a supervisão do sistema financeiro, avaliação e gestão de risco.

Para ilustrar alguns dos produtos e serviços produzidos pelo BdP, foram extraídos dois relatórios do *BPstat*: a caracterização do agregado do Quadro de Setor (Quadro 16 do Anexo) e os rácios económico financeiros europeus do setor da Construção (Quadro 17 do Anexo), com prioridade para a classe de dimensão de grandes empresas, relativo ao período de 2013 e 2014.

5.1. Quadros de Empresa e Setor, Quadros de Setor e a Série Longa dos Quadros de Setor

Os Quadros de Empresa e Setor (QES) encontram-se disponíveis a partir de 2010 e as empresas podem consultar a sua informação anual, mediante autenticação prévia, assim como também podem consultar a informação do agregado do setor e da classe de dimensão a que pertencem.

Os indicadores dos quadros do setor estão organizados de acordo com a natureza da informação e da análise que se pretende realizar. Estes quadros englobam a caracterização do agregado, indicadores de síntese, balanço, demonstração de resultados, fluxos de caixa e rácios económico-financeiros e europeus.

A caracterização do agregado reflete o número de empresas do agregado em determinado período, enquanto a representatividade é refletida pela percentagem das empresas do agregado e permite constatar em que intervalo percentual se situam as empresas, pessoas ao serviço, volume de negócios e acontecimentos marcantes, como cisões e fusões de empresas (Quadro 16).

Adicionalmente pode ser apresentada a distribuição das empresas do agregado em percentagem, de acordo com a localização da sede, estabelecimentos das empresas, natureza jurídica e maturidade das SNF.

Os indicadores de síntese avaliam a situação do agregado do setor e a classe de dimensão que é atribuída em função do volume de negócios. Os valores são apresentados através da média do agregado.

O balanço fornece indicadores sobre a situação patrimonial da empresa à data de fecho das contas, sendo que a síntese está apresentada de acordo com os normativos contabilísticos em vigor, procurando conciliar os distintos regimes de relato financeiro.

A demonstração de resultados fornece a panorâmica geral sobre a atividade desenvolvida pelas empresas durante o ano económico, assim como detalhes sobre os rendimentos e os gastos realizados no período anual. Inclui ainda indicadores adicionais que refletem a relação da empresa com o exterior (isto é, as trocas comerciais), bem como o valor acrescentado bruto e o resultado de exploração (que não incluiu os resultados financeiros e fiscais).

Os fluxos de caixa refletem os fluxos financeiros gerados ou consumidos, pelas empresas. De realçar, no entanto, que o modelo usado não apresenta correspondência com o normativo contabilístico nacional, associado às limitações das empresas não sujeitas a este reporte como é o caso das micro empresas.

Com base na demonstração de resultados e no balanço, são calculados os rácios económicos e financeiros para realizar a análise da liquidez, da estrutura financeira, do financiamento, da rentabilidade, do risco, de atividade, técnicos e de repartição de rendimentos do setor numa base anual. Os rácios económicos e financeiros são calculados e apresentados com base na média aparada e na média do agregado. A Série Longa dos Quadros do Setor (SLQS) divulga um conjunto de indicadores mais restritos, em consonância com as normas contabilísticas em vigor e que garantem um elevado nível de comparabilidade dos conceitos nas séries existentes, apresenta os indicadores comuns ao balanço e à demonstração de resultados e os rácios económico-financeiros.

5.2. Boletim Estatístico

O Boletim Estatístico é uma publicação mensal *bilingue* (português/inglês), através da qual o BdP divulga informação estatística compilada no âmbito das suas atribuições, nomeadamente: estatísticas monetárias, financeiras, cambiais, da balança de pagamentos, da posição de investimento internacional e das contas nacionais financeiras. O Boletim inclui ainda capítulos específicos sobre

estatísticas das finanças públicas, estatísticas das empresas não financeiras da Central de Balanços, estatísticas de emissão monetária e estatísticas gerais da economia portuguesa.

5.3. Serviços

O BdP dispõe de serviços de divulgação da informação estatística, nomeadamente o BPstat - Estatísticas Online e, mais recentemente, as Notas de Informação Estatística que são publicadas de forma regular através do portal da instituição.

O BPstat - Estatísticas Online do BdP corresponde a um sistema de acesso livre e disponível online através da Base de Dados Estatística Interativa do BdP. Esta ferramenta possibilita a difusão de estatísticas relevantes sobre a economia portuguesa, estando também disponível em versão *bilingue* (*português/inglês*). O seu objetivo principal consiste em facultar um acesso fácil e célere às séries estatísticas produzidas pelo BdP, assim como o acesso às principais estatísticas e indicadores económicos da responsabilidade de outras instituições. Além disso, este serviço disponibiliza um conjunto alargado de informação estatística (dados e metainformação) e diversas funcionalidades que permitem aos utilizadores personalizar e utilizar este sistema de forma interativa.

5.4. Portugal no Grupo BACH

Portugal integra o *Bank for the Accounts of Companies Harmonized* (BACH) e é inclusive um dos países de referência no reporte das estatísticas das SNF neste grupo.

A base de dados do BACH foi criada pelo ECCBSO (*European Committee of Central Balance-Sheet Data Offices*) e contém informação contabilística anual harmonizada sobre as estruturas do balanço e da demonstração de resultados das empresas não financeiras de 11 países europeus: Alemanha, Áustria, Bélgica, Eslováquia, Espanha, França, Holanda, Itália, Portugal, Polónia e República Checa. Outros quatro países europeus que integram o BACH são a Dinamarca, o Luxemburgo, a Roménia e a Turquia.

O ECCBSO é um órgão consultivo criado em 1987 por um grupo de bancos centrais nacionais europeus, cujo objetivo é melhorar a análise das SNF, promovendo o trabalho realizado pelas

diferentes centrais de balanços nacionais entre os membros do Sistema Europeu de Bancos Centrais (SEBC) e explorar as sinergias com o Banco Central Europeu (BCE), o Eurostat e os institutos nacionais de estatística, no âmbito dos dados de balanço das SNF relevantes para as funções do SEBC e do Eurossistema, bem como para o Sistema Estatístico Europeu (SEE). A presidência do ECCBSO esta a cargo do Banco de Portugal, desde novembro de 2013, até novembro 2016.

A título exemplificativo pode ser consultado em Anexo o Quadro 17 onde são apresentados os rácios setoriais europeus (a partir de 2006) obtidos da base de dados BACH para a França e Portugal.

A importância da base de dados BACH é reconhecida pela Autoridade Bancária Europeia visto que "a base de dados BACH é a fonte mais confiável de dados disponíveis que contém o balanço e a demonstração de resultados das empresas europeias." (EBA, 2012).

Uma vez que Portugal é um dos países que integra o BACH, optou-se pelo estudo da metodologia utilizada pelo BdP na realização do caso prático.

III. Caso Prático: Aplicação da Produção de Estatísticas das Empresas Não Financeiras à realidade de Angola

O estudo acerca da Produção de Estatísticas das Empresas Não Financeiras, aplicado à realidade de Angola, foi desenvolvido no âmbito do estágio curricular realizado na Central de Balanços do Departamento de Estatística do Banco de Portugal, para a conclusão do Mestrado em Matemática e Aplicações. O estágio e o estudo que dele resultou apenas foram possíveis no âmbito do acordo de cooperação institucional celebrado entre os bancos centrais, de Portugal e Angola. A cooperação entre os bancos centrais dos dois países decorre já a outros níveis, nomeadamente o da formação bancária, uma vez que Portugal apoia Angola na formação e capacitação técnica dos seus quadros.

À semelhança de Angola, a cooperação com o BdP tem beneficiado os outros países africanos de língua oficial portuguesa (PALOP), tais como Cabo Verde, Guiné-Bissau, Moçambique, São Tomé e Príncipe e Guiné Equatorial. Ao nível da formação bancária, o BdP tem ministrado cursos em gestão bancária, técnicos e administrativos, aos funcionários dos Bancos Centrais dos PALOP, em coordenação com o Fundo Monetário Internacional, com vista a uma adequação de realidade dos bancos centrais em conformidade com as práticas internacionais.

O caso prático que se apresenta corresponde a uma adequação da metodologia implementada pelo BdP na produção de estatísticas das SNF ao caso angolano, tendo em conta a realidade da estrutura dos dados atualmente fornecidos pelas SNF. Pretende-se, deste modo, contribuir com uma proposta de implementação de uma nova dinâmica de publicação de resultados que se traduza em mais-valias para a produção de estatísticas, fornecendo melhores instrumentos para análise económica e financeira da atividade empresarial.

Muito embora a metodologia utilizada pelo BdP contemple procedimentos para o tratamento da informação trimestral e conciliação com os dados anuais, não é possível aplicar esta metodologia na íntegra à realidade de Angola, uma vez que as empresas nacionais reportam a sua informação contabilística numa base anual. Contudo este facto não invalidou o estudo integral da metodologia e das práticas utilizadas pelo BdP.

6. Enquadramento Macroeconómico

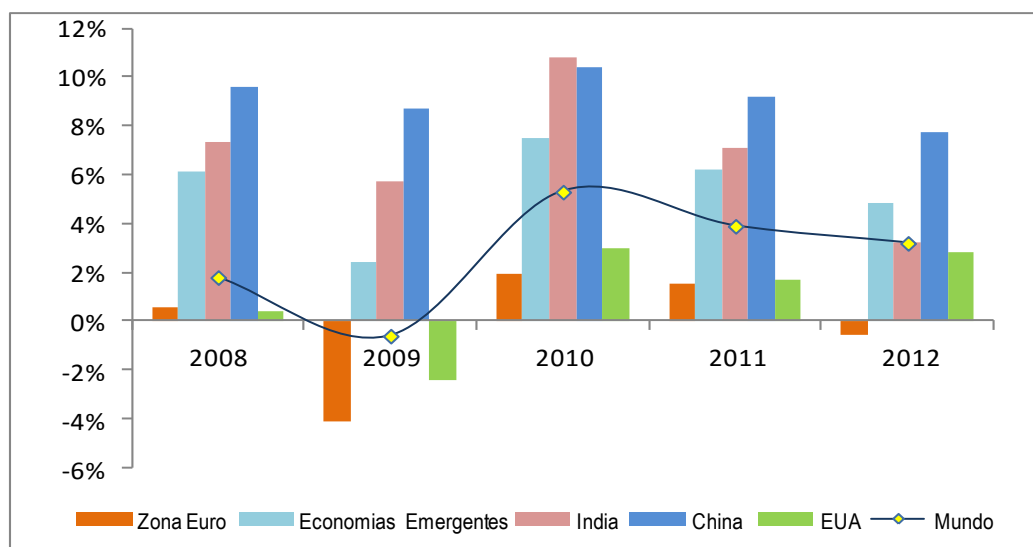
Numa primeira fase foi efetuado um enquadramento da economia internacional no ano de 2012, assim como uma análise da economia de Angola no mesmo período.

6.1. Economia Internacional

A economia mundial em 2012 foi marcada pela crise da dívida soberana de alguns países da zona Euro, nomeadamente Grécia, Espanha, a Irlanda, Portugal e Itália, o que provocou uma crise de confiança nos mercados financeiros.

O Gráfico 1 representa a evolução do PIB mundial e do PIB das economias de referência (China, EUA, zona Euro, economias emergentes e Índia). O PIB da China foi o mais elevado ao longo de 2008 a 2012. A partir de 2010 todas as economias sofreram uma queda no PIB, exceto a dos EUA.

Gráfico 1. Gráfico de linhas sobre a evolução do PIB Mundial, entre 2008 e 2012 (%)



Fonte: *World Economic Outlook*, 2013

A taxa média de crescimento a nível mundial agravou-se ao passar de 3,90% em 2011 para 3,20% em 2012. Relativamente à Zona Euro, esta sofreu uma contração de 1,00 ponto percentual (ao passar de 1,40% em 2011 para -0,40% em 2012) enquanto as economias emergentes, EUA, China e Índia cresceram 4,80%, 2,80%, 7,70% e 3,20% respetivamente (Gráfico 1).

A desaceleração das economias foi influenciada por fatores como o elevado nível de desemprego, a fragilidade do setor financeiro e a austeridade fiscal que deprimiram a procura,

enquanto o crescimento de outros países foi possível graças às políticas monetárias adaptáveis dos bancos centrais que lançaram novos programas de compra de títulos e mantiveram as taxas de juros baixas.

A nível mundial, a taxa de desemprego registou uma tendência negativa crescente, atingindo os 8,00% em 2012. Esta foi agravada principalmente pela zona Euro, onde se registou, em 2012, um aumento de 12,00% no desemprego. Contrariamente, nos EUA, Canadá e Japão a taxa de desemprego decresceu.

A taxa de inflação nas economias desenvolvidas da zona da Europa, em 2012, manteve a tendência decrescente, tendo atingido os 2,40%. Nos países emergentes e em desenvolvimento, a taxa de inflação sofreu um declínio mas não tão acentuado, de 1,10 ponto percentual ao passar de 7,20% em 2011 para 6,1% em 2012, devido a uma melhoria na credibilidade da política monetária, associado ao crescimento económico. Nas economias desenvolvidas, a taxa de inflação também decresceu, registando uma queda menos acentuada de 0,70 pontos percentuais ao atingir 2,00% em 2012 contra 2,70% de 2011.

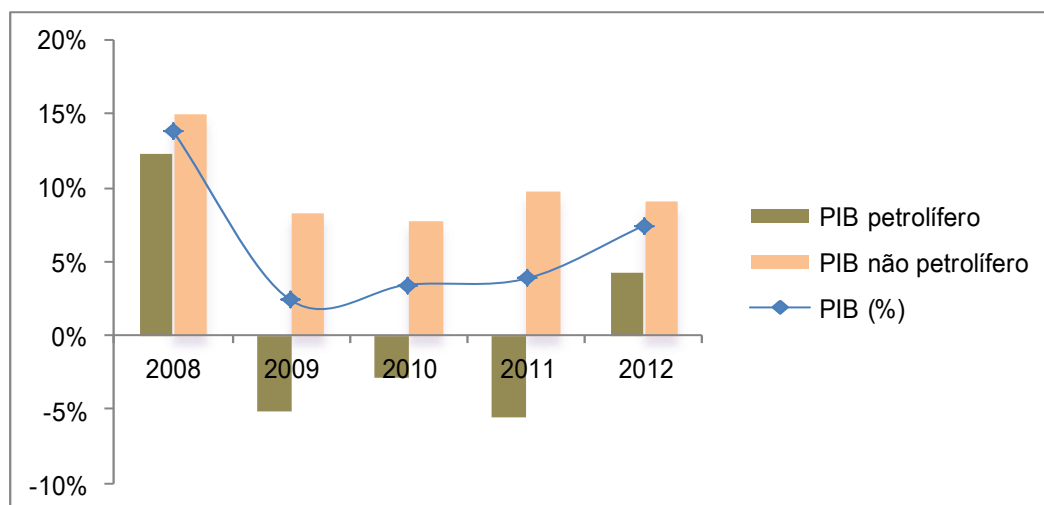
A variação da taxa de inflação pode estar relacionada com fatores como o preço do petróleo, o aumento do preço dos alimentos, a redução do PIB, a depressão do mercado de trabalho e a credibilidade da Política Macroeconómica.

6.2. Economia Angolana

Em 2009, o Governo de Angola beneficiou de um financiamento de 1,4 mil milhões de dólares, no âmbito de um acordo *Standby* assinado com o FMI. A sua implementação decorreu entre 2009 e 2012, com os objetivos de aliviar as pressões de liquidez, restabelecer a confiança do mercado, melhorar a situação macroeconómica, realizar reformas estruturais, e ainda estabilizar o crescimento do setor não petrolífero.

Pelo Gráfico 2 pode observar-se que a crise de 2008-2009 influenciou substancialmente o crescimento da economia Angolana, pois o PIB sofreu um decréscimo acentuado passando de 11,2% (em 2008) para 2,1% (em 2009).

Gráfico 2. Evolução do PIB, do PIB petrolífero e do PIB não petrolífero angolano (%)



Fonte: Ministério do Planeamento/Ministério das Finanças/BNA, 2013

Observou-se a partir de 2010, uma recuperação do PIB, tendo este tingido em 2012 uma taxa de 7,45%, enquanto que em 2011 esta era de 3,86%. O desempenho da economia em 2012 está associado ao aumento da produção de petróleo. Conforme ilustra o Gráfico 2 nos anos de 2009, 2010 e 2011 o PIB petrolífero contribuiu negativamente com -5,09%, -2,87% e -5,57% respetivamente, enquanto que o não petrolífero influenciou positivamente ao atingir 8,31%, 7,80% e 9,74%, respetivamente, para o crescimento da economia angolana. Dentro do setor não petrolífero distinguem-se os serviços mercantis com uma contribuição de 21,68%, a agricultura e pescas com 9,51%, a construção com 8,39% e a indústria transformadora com 6,47%.

No final do ano de 2012, a inflação situou-se em 9,00%. O crescimento das receitas petrolíferas provocaram um impacto positivo no saldo orçamental global ao atingir um *superavit* de 8,50% do PIB. A 31 de Dezembro de 2012, as reservas internacionais subiram para o equivalente a 7,3 meses de importações.

7. Enquadramento Legal em Angola

O Decreto nº82/01, de 16 de novembro, criou o Plano Geral de Contabilidade de Angola com vista a harmonizar as práticas nacionais com as internacionais, obrigatoriamente aplicável às Empresas Públicas e Sociedades Comerciais que atuam ou que possuem sede em Angola. Assim, as Empresas Públicas e as Sociedades Comerciais reportam o Relatório e Contas, numa base anual e em suporte de papel, ao Ministério das Finanças para efeitos fiscais. Desta obrigação estão isentos os particulares e as empresas com planos de contabilidade próprios, nomeadamente a área dos Seguros e a Banca.

O subsetor das SNF abrange o conjunto das sociedades e quase-sociedades não financeiras que não são controladas pelas administrações públicas ou pelas unidades institucionais não residentes. Este subsetor engloba igualmente todas as sociedades e quase-sociedades de investimento estrangeiro direto que não se classificam no subsetor das sociedades não financeiras de controlo estrangeiro, bem como as instituições sem fins lucrativos.

De acordo com a Lei n.º 15/96, de 27 de setembro, Lei de Bases do Sistema Nacional de Estatística, a informação sobre as SNF é publicada de forma agregada, não sendo possível o acesso e a publicação de dados micro das empresas.

De acordo com o da Lei do Banco Nacional de Angola (Lei n.º 16/10, de 15 de julho), para além da condução, execução, acompanhamento e controlo das políticas monetária, financeira, cambial e de crédito no âmbito da política económica do Poder Executivo, compete ao Banco Nacional de Angola (BNA): “garantir e assegurar um sistema de informação, compilação e tratamento das estatísticas monetárias, financeiras e cambiais e demais documentação, nos domínios da sua atividade por forma a servir como instrumento eficiente de coordenação, gestão e controlo; elaborar e manter atualizado o registo completo da dívida externa do País, assim como efetuar a sua gestão; elaborar a balança de pagamentos externos do País” (n.º2 do artigo 16 da Lei n.º 16/10, de 15 de julho).

Relativamente à classificação das empresas, foi publicada a Lei das Micro, Pequenas e Médias Empresas (Lei n.º 30/11, de 13 de setembro) que considera critérios distintos nesta classificação: o número de trabalhadores efetivos e o volume anual de faturação, sendo este último o requisito prevalecente em caso de dúvida entre duas categorias de classificação.

De realçar que existe uma diferença relativamente aos critérios para classificar a dimensão da empresa, comparativamente ao Banco de Portugal, não é considerado o total de ativo para determinar a dimensão da empresa.

A referida Lei considera micro empresas aquelas que possuem uma faturação bruta anual inferior em kwanzas ao equivalente a USD 250.000,00 (Duzentos e Cinquenta mil dólares americanos) e empreguem até 10 trabalhadores. As pequenas empresas possuem uma faturação bruta anual em kwanzas superior ao equivalente a USD 250.000,00 (Duzentos e Cinquenta mil dólares americanos) e igual ou inferior a USD 3.000.000,00 (Três milhões de dólares americanos) e empreguem mais de 10 trabalhadores e até 100. As médias empresas serão as que tenham uma faturação bruta anual em kwanzas superior ao equivalente a USD 3.000.000,00 (Três milhões de dólares americanos) e igual ou inferior a USD 10.000.000,00 (Dez milhões de dólares americanos) e empreguem mais de 100 e até 200 trabalhadores.

Tradicionalmente, as Micro, Pequenas e Médias Empresas (MPME) têm sido dos principais instrumentos de sustentação das economias modernas, incluindo as dos países mais desenvolvidos, não apenas por participarem na redução do desemprego, mas também por se ajustarem às necessidades das comunidades e, com isso, contribuírem, significativamente para a redução da pobreza. Em Angola, a adoção e implementação de uma estratégia de fomento das MPME recomenda a adoção prévia de um amplo programa de simplificação das práticas administrativas, de regulamentação e de facilitação do acesso aos mercados e a novas oportunidades de negócios, bem como a formalização de parcerias visando o desenvolvimento de novos produtos e serviços, em geral.

Assim, para a formulação das políticas e programas dirigidos às empresas não financeiras, é importante proceder à análise da evolução dos indicadores financeiros das mesmas, por forma a detetar e dar resposta às necessidades do setor.

8. Descrição do Problema em Estudo

O problema da produção das estatísticas das SNF reveste-se de extrema importância para Angola, visto que, à semelhança de uma grande parte dos países da África Austral e dos PALOP, o país não dispõe de uma Central de Balanços para gerar estatísticas das SNF.

Em Portugal, a CB é a área responsável por compilar e publicar as estatísticas das SNF, a partir de micro dados contabilísticos obtidos a partir do ITENF (de base trimestral) e do IES (de base anual) em suporte eletrónico e abrangendo todos os setores de atividade das SNF. O ITENF permite recolher uma amostra das SNF bem como extrapolar e conciliar dados para o universo das SNF.

Os resultados disponibilizados pela CB permitem complementar a informação estatística que é produzida e divulgada para as SNF a partir de outros sistemas de informação, como por exemplo no âmbito das estatísticas monetárias e financeiras, da balança de pagamentos e posição de investimento internacional, e das contas financeiras.

As estatísticas das sociedades não financeiras da CB refletem a partir de informação de base contabilística, a evolução da situação económica e financeira do conjunto de empresas que integra o setor das SNF em Portugal (como definido no sistema de contas nacionais), com exceção das pertencentes à secção A da Classificação Portuguesa das Atividades Económicas, Revisão 3 (CAE – Rev.3): Agricultura, produção animal, caça, floresta e pesca. (BdP, 2013).

A obtenção deste tipo de indicadores da atividade económica e financeira sobre o tecido empresarial é do interesse de Angola, bem como dos demais países africanos, pelo que se considera pertinente a implementação de uma base de dados integrada com a informação contabilística das SNF.

Para o BNA a implementação de uma Central de Balanços, torna possível a elaboração de estatísticas das SNF, a realização de estudos setoriais e o cruzamento da informação contabilística com o Departamento de Central de Risco, no que concerne ao endividamento privado.

Poderá também fornecer informação ao Setor Externo, nomeadamente a dívida privada e os lucros e dividendos do setor petrolífero. Esta informação é atualmente objeto de tratamento manual, após disponibilização do relatório e contas na base de dados micro pelo Ministério das Finanças, necessários para a compilação da Balança de Pagamentos e o Departamento de Controlo Cambial que poderá utilizar os dados sobre a situação económica e financeira das empresas aquando do licenciamento das operações de invisíveis e de capital.

Atualmente, Angola tem implementado o Inquérito Anual Harmonizado às Empresas (IAHE) através do qual recolhe, numa base anual, dados económico-financeiros, em particular a demonstração de resultados das empresas não financeiras. Encontra-se excluído do âmbito de aplicação do IAHE o setor da Agricultura e Pescas (CAE-Rev.2, A). O IAHE recolhe informação detalhada sobre os serviços de terceiros, investimentos realizados pelas empresas, o volume de ativo e dados qualitativos, como a localização e os contatos das empresas.

O IAHE tem como objetivos gerais e específicos: disponibilizar informação estatística detalhada e de qualidade para as Contas Nacionais; fornecer informação detalhada para a atualização do Ficheiro de Unidades Empresariais (Empresas e Estabelecimentos); analisar o desenvolvimento regional e nacional das empresas; disponibilizar aos utilizadores, estatísticas detalhadas sobre a atividade

empresarial; atualizar o conjunto dos indicadores sobre a atividade empresarial, bem como fornecer elementos para a implementação de uma central de balanços (INE Angola, 2015).

O último IAHE foi publicado em 2015 com dados referentes ao período de 2009 a 2012. O documento analisa apenas dados relativos ao número total de empresas com atividade em Angola, no período considerado, segundo a distribuição por Província, forma jurídica e empresas importadoras e/ou exportadoras (INE Angola, 2015). O mesmo documento reúne, ainda, dados relativos à evolução do montante do volume de negócios (Kwanzas) e número de pessoas ao serviço.

O referido documento não inclui a totalidade das demonstrações financeiras previstas no Plano Geral de Contabilidade (PGC) (designadamente, o balanço, os fluxos de caixa das empresas, as notas às contas das empresas não financeiras, entre outros). Desta forma, não é possível realizar a análise económica financeira, nem a sua avaliação patrimonial, uma vez que não possui informação em detalhe suficiente para o efeito. O inquérito é de âmbito nacional, de natureza económica e financeira, constitui a fonte de informação estatística das empresas, tendo como objetivo o conhecimento sobre a sua estrutura, evolução da atividade empresarial, bem como produção de informação de base para as contas nacionais, usuários e suportes para o planeamento e tomada de decisão para o Executivo (INE Angola, 2015).’

O caso prático em estudo teve, por isso, como objetivo a adequação da aplicabilidade da metodologia em uso no Banco de Portugal para a compilação das estatísticas das SNF em Angola, tomando em consideração a especificidade de reporte dos dados efetuado na base anual. Por outro lado, pretendeu-se com a elaboração do quadro de setor efetuar a análise económico-financeira do agregado em estudo, nomeadamente um grupo de empresas portuguesas de construção que atuaram em Angola, de 2010 a 2012.

9. Recolha da Informação

Para concretizar os objetivos propostos foi necessária a constituição de uma base de análise com os dados referentes ao agregado. Associado ao fator de limitação do acesso aos dados micro das SNF (Lei n.º 15/96, de 27 de setembro), o caso prático foi restrito ao setor da construção. A constituição do agregado atendeu assim primeiramente ao critério da classificação da atividade económica (CAE-Rev.2, “F-Construção”) e reuniu dados contabilísticos, relativos a 2010, 2011 e 2012, de onze empresas do setor da construção a operar em Angola das quais cinco se encontravam no *ranking* das 10 maiores empresas em Portugal.

De acordo com a informação disponível, 81,82% das empresas do agregado eram micro empresas (n=9) e as restantes 18,18% eram pequenas e médias empresas (n=2). A dimensão das empresas que constituíram o agregado foi assumida de acordo com o critério do volume de faturação bruta anual das empresas (Lei n.º 30/11, de 13 de setembro).

Como não foi possível obter os dados informatizados (os dados foram fornecidos em suporte de papel) optou-se por utilizar o *Microsoft Excel* do *Office* para a compilação da base de dados referente à demonstração de resultados e ao balanço das empresas em estudo, de acordo com a estrutura definida pelo Plano Geral de Contabilidade de Angola (Decreto nº82/01, de 16 de novembro).

Depois de constituída a base de dados, optou-se pelo uso do *software R Statistics (R-Project)* para fazer a imputação dos dados omissos e obter as medidas de localização e dispersão dos rácios. Posteriormente, a obtenção dos rácios económico-financeiros e dos quadros de setor foi feita com o *Microsoft Excel*.

Na avaliação prévia da informação recolhida sobre as empresas do agregado, para os dados referentes a 2010 e 2011 observou-se uma taxa de não resposta de 18,18% (o que correspondeu a três empresas), enquanto para os dados de 2012 a taxa de não resposta foi de 9,09% (o que correspondeu a uma única empresa).

Na compilação de dados económicos e financeiros é frequente os investigadores e analistas se deparem com dados omissos ou ausentes (*missing data*). Os dados *omissos* afetam não apenas a sua consistência, como também a qualidade das estatísticas produzidas. Por isso, determinar a abordagem analítica adequada ao tratamento dos dados omissos requer métodos e procedimentos específicos, uma vez que estes têm impactos nos resultados e nas ilações a retirar sobre o conjunto dos dados. Uma vez que a falta de dados na amostra pode gerar viés e ineficiência nas estimativas, não podemos deixar de considera-los nas análises (Nunes, Klück & Fachel, 2010).

O desenvolvimento dos métodos estatísticos de imputação de dados têm sido uma área de pesquisa bastante ativa nas últimas décadas: Donald Rubin (1987) desenvolveu técnicas de imputação múltipla na década de 80, tendo-se seguido vários outros investigadores no estudo de técnicas estatísticas que envolvem a imputação dos dados omissos, no sentido de os substituir por valores plausíveis (Little, 1992; Zhang, 2003).

As primeiras técnicas de imputação desenvolvidas eram mais simples, dado que os dados faltantes eram substituídos pela média, pela mediana, ou estimados por interpolação ou regressão linear (Nunes, Klück & Fachel, 2010). São as chamadas imputações únicas, onde o valor ausente é substituído apenas uma vez, obtendo-se a base de dados completa e preparada para a análise com as

técnicas tradicionais. Embora sejam utilizados devido à sua facilidade de aplicação, as técnicas de imputação única, apresentam algumas desvantagens como a subestimação da variabilidade da variável imputada e a dificuldade de levar em consideração a variabilidade que possa existir entre diferentes imputações (Nunes, Klück & Fachel, 2010). Por isso, seria recomendável considerar na análise a incerteza adicional associada ao processo de imputação, uma vez que o valor imputado não é o verdadeiro valor da variável.

No caso prático em estudo, o problema da não resposta exigiu o tratamento estatístico segundo procedimentos específicos de imputação de dados, através da utilização de informação na ótica contabilística, nomeadamente, a demonstração de resultados e o balanço das empresas portuguesas que constituíram o estrato. O tratamento não resposta foi realizado através do software *R*, testado em duas variantes uma com base na imputação simples pela média do agregado e outra pelo método Cold-Deck (Little, R. J. A., 1988).

Para salvaguardar a consistência da informação contabilística e implementar alguns procedimentos de controlo de qualidade da informação, na 1ª etapa de compilação das estatísticas das SNF, ao nível do balanço e da demonstração de resultados, foram incluídas algumas validações automáticas, nomeadamente, controlos para validar a igualdade entre as somas parcelares e as somas dos subtotais por classe, soma dos subtotais com a soma do total da conta, validação de variáveis que não podem assumir negativos, equilíbrio entre o ativo e o passivo, bem como alertas de sinalização de variações absolutas e relativas, que excedam os limites estabelecidos, (valor absoluto na ordem de Quinze Mil milhões de Kwanzas).

Para a imputação dos dados omissos foi utilizada a informação do período 2010-2012, enquanto a análise económica financeira foi realizada para o período de 2012, comparativamente a 2011.

10. Tratamento da Informação Anual

Nesta secção procede-se à descrição dos dois procedimentos considerados no tratamento dos valores omissos na base de dados, os métodos de imputação simples pela média e Cold-Deck, cuja seleção é efetuada no ponto seguinte.

10.1. Tratamento da Não Resposta

Com o tratamento de não resposta pretendemos obter informação para os dados omissos ao nível do balanço e da demonstração de resultados. Na imputação de dados omissos foram testadas duas variantes, a variante I que se refere à imputação pela média do estrato das empresas com dados completos, e a variante II que imputa os valores em falta pelo método Cold-Deck.

O Quadro 8 apresenta os pressupostos comuns usados nas duas variantes para o tratamento não resposta (através do método de imputação pela média e através do método Cold-Deck), nomeadamente o número de empresas, a estrutura do agregado e a taxa de não resposta.

Quadro 8. Pressupostos Comuns para a Variante I e Variante II

	2010		2011		2012	
	ME	PE	ME	PE	ME	PE
Número de Empresas	9	2	9	2	9	2
Total de Empresas	11		11		11	
Nível de não resposta por dimensão da empresa	18,18	0,00	18,18	0,00	9,09	0,00

Legenda: ME – Micro Empresa; PE – Pequena Empresa.

No Quadro 9, são apresentados os pressupostos específicos para cada variante (métodos), que a seguir se descreve.

Quadro 9. Pressupostos Específicos da Variante I e Variante II

	VARIANTE I	VARIANTE II
DR	Média da variação do estrato DR	Média da variação de Volume de Negócios
Balanço	Média da variação do estrato do Balanço	Média da variação de Ativo

Legenda: DR - Demonstração de Resultados

Na variante I – (Imputação pela Média) para o balanço calculou-se, a média para as variáveis observadas, com base nos dados completos das empresas do mesmo estrato/dimensão, para determinado período, sem haver relação com os dados históricos das empresas com dados omissos. Este procedimento é semelhante para a demonstração de resultados.

A variante II – (Cold-Deck) tem em conta os dados históricos da empresa que apresenta valores em falta (ou omissos), num determinado período. No caso da demonstração de resultados, sobre os

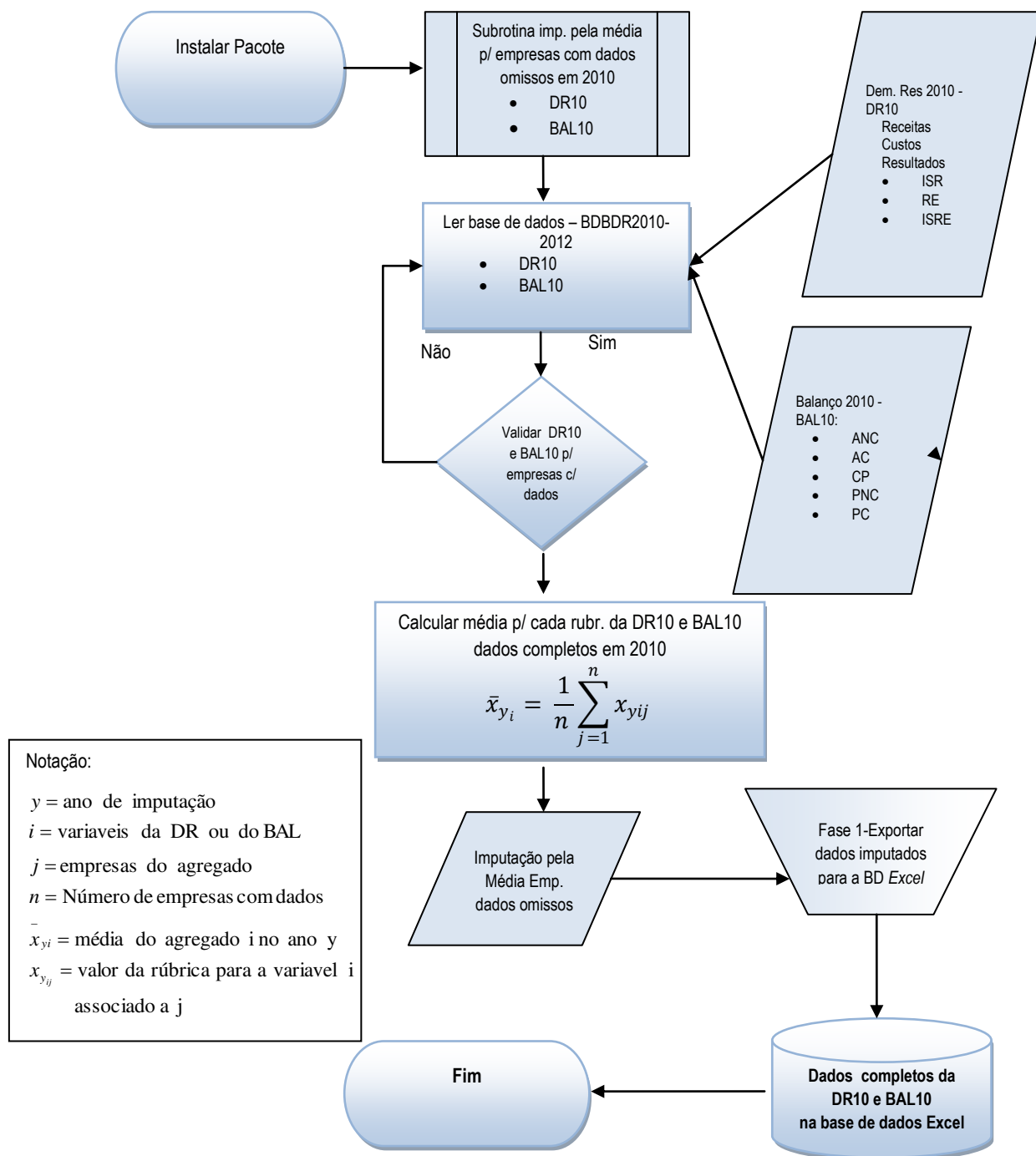
dados históricos, recaí o efeito da média de variação do volume de negócios das empresas com informação completa do mesmo estrato e dimensão para o período em análise. Em relação ao balanço, foi utilizada a taxa média de variação do total de ativo do agregado.

10.1.1. Imputação pelo Método da Média – Variante I

A imputação pela média no estrato é conhecido como método de imputação simples e é usado com mais frequência pela sua facilidade de implementação. Este método apresenta algumas desvantagens como a subestimação da variabilidade da variável imputada e a dificuldade de levar em consideração a variabilidade que possa existir entre diferentes imputações.

No caso prático, foi utilizada apenas a base de dados criada em *Excel*, devido à inexistência de bases de dados alternativas. Começa-se por ilustrar o funcionamento da variante I, mediante o fluxograma da Figura 8.

Figura 8. Imputação pela Média – Variante I



Nota: Demonstração de Resultados → ISR - Imposto Sobre Rendimento, RE- Resultados do Exercício, ISRE- Imposto sobre resultados extraordinário ISRE- Imposto sobre resultados extraordinário;
Balanco → ANC – Ativo não corrente, AC – Ativo corrente, CP- Capital próprio, PNC- Passivo , PC- Passivo Corrente.

10.1.2. Imputação pelo Método Cold-Deck – Variante II

O método Cold-Deck pressupõe a utilização de dados históricos da empresa existentes numa base de dados, ou numa base de dados alternativa, podendo ser utilizados dados de período anteriores ou posteriores ao ano a que se refere a imputação.

Para implementação do método Cold-deck as empresas do agregado foram estratificadas com base no setor de atividade e dimensão das empresas. No caso prático, foi efetuada a imputação de dados omissos para a empresa 1, classificada como micro empresa, tendo em conta o seu volume de negócios, inferior a 250 mil dólares americanos. Para uma melhor ilustração do método é apresentado um exemplo com apenas três empresas (empresa 1, empresa 2, empresa 8), tendo as empresas 1 e 8 dados completos.

Etapas 1

Para as empresas que não cumprem com a obrigação de reportar a informação, recomenda-se a realização do tratamento não resposta. No Quadro 10 é apresentado o balanço da empresa 1 (EMP1) relativo aos exercícios económicos de 2010 a 2012, podendo se constatar que essa empresa não respondeu ao inquérito do ano de 2010.

Na etapa 1, é selecionado um ano anterior ou posterior, em que a EMP1 possui dados históricos, bem como as empresas do mesmo estrato/dimensão que possuem dados completos. Para o efeito selecionou-se o ano de 2012, a empresa 2 (EMP2) e a empresa 8 (EMP8) que pertencem ao estrato da construção, dimensão micro empresas, e volume de negócios inferior ao equivalente a USD 250.000,00.

Etapas 2

O estrato selecionado é integrado por nove microempresas consideradas na análise. No entanto, no Quadro 11 são apresentadas apenas duas empresas (EMP2, EMP8) que serão utilizadas para ilustrar a aplicação do método.

Na etapa 2, calcula-se o total de ativo para a EMP2 e EMP8 para os anos de 2010 e 2012 (Quadro 12), com vista a obtenção da taxa de variação da variável do total de ativos

A taxa de variação anual do total de ativo para as empresas com dados completos nos dois anos é calculada usando a expressão:

$$\nabla TA = \frac{\sum_{i=1}^n TA_{2012} - \sum_{i=1}^n TA_{2010}}{\sum_{i=1}^n TA_{2010}} = \frac{\sum_{i=1}^n (TA_{2012} - TA_{2010})}{\sum_{i=1}^n TA_{2010}}$$

em que $\sum_{i=1}^n TA_{2010}$ e $\sum_{i=1}^n TA_{2012}$ correspondem à soma do total do ativo de todas empresas com dados em 2010 e 2012, respetivamente, e n representa o número total de empresas com dados completos.

Em seguida calcula-se a taxa média de variação do total de ativo, que se obtém mediante a expressão:

$$TM(\nabla TA) = Média[\nabla TA] = \frac{1}{n} \sum_{i=1}^n \nabla TA$$

em que $\sum_{i=1}^n \nabla TA$ corresponde à soma das taxas de variação anual de todas as empresas, e n ao número total de empresas com dados completos.

Etapa 3

Por último na etapa 3, é efetuada a imputação de dados omissos à EMP1, para cada variável do seu balanço de 2010, sobre as quais recai o efeito da taxa média de variação da variável total de ativos calculada para o estrato-dimensão.

A imputação das variáveis omissas, no caso da EMP1 e no ano de 2010, pode ser feita através da seguinte expressão:

$$X_{2010}(Empresa\ 1) = \frac{X_{2012}(Empresa\ 1)}{TM(\nabla TA) + 1}$$

em que $X_{2010}(Empresa\ 1)$ corresponde o dado a ser imputado à empresa 1 em 2010.

De salientar que este procedimento também é válido para a imputação de dados omissos da demonstração de resultado, sendo utilizada no caso a variável volume de negócios.

Quadro 10. Balanço da EMP1 (Em Milhões de Kwanzas)

		BALANÇO			
		ATIVO			
			2010	2011	2012
		Notas	EMP1BAL2010	EMP1BAL2010	EMP1BAL2011
ATIVO	Ativos não correntes				
	Imobilizações corpóreas	4		942,00	729,14
	Imobilizações incorpóreas	5		0,00	0,00
	Investimento em subsidiárias e associadas	6		0,25	0,00
	Outros ativos financeiros	7		0,00	0,00
	Outros ativos não correntes	9		0,00	0,00
	Total do ativo não corrente		0,00	943,07	729,14
	Ativos correntes				
	Existências	8		82, 73	77, 23
	Contas a receber	9		3.949, 28	5.504,08
	Disponibilidades	10		488, 09	31, 84
	Outros ativos correntes	11		0,00	0,00
	Total do ativo corrente		0,00	4.520,09	5.613,15
	Total do Ativo		0,00	5.463,17	6.342, 29
CAPITAL PROPRIO E PASSIVO					
CAPITAL PROPRIO E PASSIVO	Capital Próprio				
	Capital	12		105,79	105,79
	Reservas	13		94,82	94,82
	Resultados transitados	14		0,00	0,00
	Resultado do exercício	14		1.359,23	2.247, 70
	Total do capital próprio		0,00	1.559,84	2.448,30
	Passivo não corrente				
	Empréstimos de médio e longo prazo	15		255,89	95,22
	Impostos Deferidos	16		0,00	0,00
	Provisões para pensões	17		0,00	0,00
	Provisões para outros riscos e encargos	18		147,00	147,00
	Outros passivos não correntes	19		0,00	0,00
	Total do passivo não corrente		0,00	402,89	242,22
	Passivo Corrente				
	Contas a pagar	19		2.466,61	2.879,29
	Empréstimos de curto prazo	20		733,45	73,00
	Parte corrente emp. a médio e longo prazo	15		0,00	0,00
	Outros passivos correntes	21		300,38	699,47
	Total do passivo corrente		0,00	3.500,44	3.651,76
	Total do passivo		0,00	3.903,33	3.893,99
	Total capital próprio e passivo		0,00	5.463,17	6.342,28

Fonte: Banco Nacional de Angola

Quadro 11. Balanço das EMP2 e EMP8

(Em Milhões de Kwanzas)

BALANÇO								
ATIVO								
ATIVO		Notas	2010	2011	2012	2010	2011	2012
			EMP2BAL2010	EMP2BAL2011	EMP2BAL2012	EMP8BAL2010	EMP8BAL2011	EMP8BAL2012
ATIVO	Ativos não correntes							
	Imobilizações corpóreas	4	1.486,52	1.406,20	3.299,10	1.082,86	1.563,94	1.998,19
	Imobilizações incorpóreas	5	3,38	1,94	0,35	10,21	10,21	10,96
	Investimento em subsidiárias e associadas	6	0,00					
	Outros ativos financeiros	7	0,00					
	Outros ativos não correntes	9	0,00	0,00				
	Total do ativo não corrente		1.489,90	1.408,14	3.299,45	1.093,17	1.574,15	2.009,15
	Ativos correntes							
	Existências	8	106,84	367,56	365,99	2,88	10,32	4,71
	Contas a receber	9	4.678,37	7.185,33	9.730,36	2.478,07	1.956,41	3.302,90
	Disponibilidades	10	1.001,95	3.127,31	571,70	653,50	767,86	695,20
	Outros ativos correntes	11	57,16	58,38	259,56	0,41	124,50	537,02
	Total do ativo corrente		5.844,31	10.738,58	10.927,61	3.134,86	2.859,09	4.539,84
	Total do Ativo		7.334,21	12.146,72	14.227,06	4.227,93	4.433,24	6.548,99
CAPITAL PRÓPRIO E PASSIVO								
CAPITAL PRÓPRIO E PASSIVO	Capital Próprio							
	Capital	12	198,00	198,00	198,00	3,75	3,75	3,75
	Reservas	13	39,60	39,60	39,60	1.092,85	2.213,07	2.626,07
	Resultados transitados	14	1.564,97	869,68	2.066,02	405,96		
	Resultado do exercício	14	2.738,71	3.159,14	3.184,59	1.714,26	913,00	2.994,18
	Total do capital próprio		4.541,28	4.266,42	5.488,20	3.216,82	3.129,83	5.624,01
	Passivo não corrente							
	Empréstimos de médio e longo prazo	15	0,00				0,00	
	Impostos Deferidos	16	0,00				0,00	
	Provisões para pensões	17	0,00				0,00	
	Provisões para outros riscos e encargos	18	263,62	23,15			0,00	
	Outros passivos não correntes	19					0,00	
	Total do passivo não corrente		263,62	23,15	0,00	0,00	0,00	0,00
	Passivo Corrente							
	Contas a pagar	19	2.396,95	7.724,34	8.032,35	43,97	537,88	890,84
	Empréstimos de curto prazo	20	0,00		500,00		0,00	
	Parte corrente emp. A médio e longo prazo	15	0,00				0,00	
	Outros passivos correntes	21	132,36	132,91	206,51	967,14	765,54	34,14
	Total do passivo corrente		2.529,31	7.857,26	8.738,86	1.011,11	1.303,42	924,98
	Total do passivo		2.792,94	7.880,30	8.738,86	1.011,11	1.303,42	924,98
	Total capital próprio e passivo		7.334,21	12.146,72	14.227,06	4.227,93	4.433,24	6.548,99

Fonte: Banco Nacional de Angola

Quadro 12. Total de ativo 2010-2012 (Em Milhões de Kwanzas)

BALANÇO								
ATIVO								
		Notas	2010 EMP2BAL2010	2012 EMP2BAL2012	2010 EMP8BAL2010	2012 EMP8BAL2012	2010 Total de ativo	2012 Total de ativo
ATIVO	Ativos não correntes							
	Imobilizações corpóreas	4	1.486,52	3.299,10	1.082,86	1.998,19	2.569,38	5.297,29
	Imobilizações incorpóreas	5	3,38	0,35	10,21	10,96	13,59	11,31
	Investimento em subsidiárias e associadas	6	0,00				0,00	0,00
	Outros ativos financeiros	7	0,00				0,00	0,00
	Outros ativos não correntes	9	0,00				0,00	0,00
	Total do ativo não corrente		1.489,90	3.299,45	1.093,17	2.009,15	2.582,97	5.308,60
	Ativos correntes							
	Existências	8	106,84	365,99	2,88	4,71	109,72	370,70
	Contas a receber	9	4.678,37	9.730,36	2.478,07	3.302,90	7.156,44	13.033,26
	Disponibilidades	10	1.001,95	571,70	653,50	695,20	1.655,45	1.266,90
	Outros ativos correntes	11	57,16	259,56	0,41	537,02	57,57	796,58
	Total do ativo corrente		5.844,31	10.927,61	3.134,86	4.539,84	8.979,18	15.467,45
	Total do Ativo		7.334,21	14.227,06	4.227,93	6.548,99	11.562,14	20.776,05
CAPITAL PRÓPRIO E PASSIVO								
CAPITAL PRÓPRIO E PASSIVO	Capital Próprio							
	Capital	12	198,00	198,00	3,75	3,75		
	Reservas	13	39,60	39,60	1.092,85	2.626,07		
	Resultados transitados	14	1.564,97	2.066,02	405,96			
	Resultado do exercício	14	2.738,71	3.184,59	1.714,26	2.994,18		
	Total do capital próprio		4.541,28	5.488,20	3.216,82	5.624,01		
	Passivo não corrente							
	Empréstimos de médio e longo prazo	15	0,00					
	Impostos Deferidos	16	0,00					
	Provisões para pensões	17	0,00					
	Provisões para outros riscos e encargos	18	263,62					
	Outros passivos não correntes	19						
	Total do passivo não corrente		263,62	0,00	0,00	0,00		
	Passivo Corrente							
	Contas a pagar	19	2.396,95	8.032,35	43,97	890,84		
	Empréstimos de curto prazo	20	0,00	500,00				
	Parte corrente emp. A médio e longo prazo	15	0,00					
	Outros passivos correntes	21	132,36	206,51	967,14	34,14		
	Total do passivo corrente		2.529,31	8.738,86	1.011,11	924,98		
	Total do passivo		2.792,94	8.738,86	1.011,11	924,98		
	Total capital próprio e passivo		7.334,21	14.227,06	4.227,93	6.548,99		

Fonte: Banco Nacional de Angola

10.2. Seleção do Método

No Quadro 13 são apresentados os indicadores sínteses gerados pelos dois métodos de tratamento não resposta, ou seja, a imputação pela média e a imputação pelo método Cold-Deck. Com vista a seleção do método de imputação, foram analisados os indicadores sínteses nomeadamente o Ativo, Capital Próprio e Volume de Negócios nos anos de 2011 e 2012. De salientar que a Variante I para estes indicadores apresentou valores superiores comparativamente aos obtidos pela Variante II.

Quadro 13. Indicadores sínteses da Variante I e Variante II, para os anos de 2010, 2011 e 2012

	U.M: Em Milhões de Kwanzas					
	Variante I - (Média)			Variante II - (Cold-Deck)		
	2010	2011	2012	2010	2011	2012
Ativo	185.943,78	249.520,19	278.291,26	186.777,23	240.538,99	268.213,97
Capital próprio	45.632,56	74.723,12	92.409,20	61.979,34	72.503,77	88.283,81
Volume de negócios	119.914,56	160.140,93	181.522,21	120.557,40	156.866,74	172.310,55
Total de rendimentos líquidos	132.631,06	181.715,48	196.709,60	130.928,00	180.075,54	192.955,75
Valor acrescentado bruto - VAB	132.631,06	181.715,48	196.709,60	130.928,00	180.075,54	192.955,75
Resultado líquido do período	24.776,90	31.182,23	27.773,70	25.903,75	29.771,88	28.622,12
Rendibilidade dos capitais próprios (%)	54,30	41,73	30,06	41,79	41,62	32,42

Fonte: Banco Nacional de Angola

Apesar da imputação pela média demonstrar valores superiores nos anos de 2011 e 2012, conforme ilustra o Quadro 13, no decorrer do estudo os valores omissos foram imputados usando o método Cold-Deck. Esta escolha residuiu no facto do método Cold-Deck considerar a variabilidade existente entre as diferentes observações. Note-se que na imputação pela média os dados omissos tomam o valor médio calculado para as empresas com informação completa do agregado em análise, não considerando os dados da empresa com dados omissos. Tal facto levará a subvalorização ou sobrevalorização dos valores obtidos, associado ainda ao facto de que para todas as empresas com dados omissos, são imputados os mesmos valores médios, calculados para as empresas com informação completa.

Após a imputação, obtiveram-se os rácios económico-financeiros do agregado, para gerar as medidas de localização com vista à compilação dos quadros de setor e proceder à análise económico-financeira, conforme se apresenta no ponto seguinte.

11. Análise Económico Financeira com Base no Quadro de Setor

Os rácios económico-financeiros são apresentados mediante a distribuição dos rácios por médias e quartis.

A informação do quadro de setor relativa aos indicadores (ou rácios) económico-financeiros foi utilizada na análise económico-financeira das empresas do agregado. No Quadro 14 são apresentadas medidas de localização anteriormente referidas para os indicadores económico-financeiros nos anos de 2011 e 2012 para o agregado em análise. O agregado em análise constitui o conjunto das 11 empresas (micro e pequenas empresas) que integram a amostra do setor da construção em Angola.

Quadro 14. Indicadores económico-financeiros

QUADRO DO SETOR										
Ano:										
Setor de atividade	F: Construção									
Classe de Dimensão:	Micro e pequenas empresas									
Rácios Económico-financeiros	2011					2012				
	Nº Emp.	1º Q	2º Q	3º Q	Média do agregado	Nº Emp.	1º Q	2º Q	3º Q	Média do agregado
Liquidez										
Liquidez geral (%)	11	122,18	135,46	165,82	145,79	11	119,49	134,37	294,15	153,00
Liquidez reduzida (%)	11	107,51	126,77	156,97	144,59	11	110,56	131,15	283,74	152,46
Estrutura financeira										
Autonomia financeira (%)	11	21,36	28,81	34,23	36,46	11	21,91	31,74	38,81	38,60
Solvabilidade geral (%)	11	27,44	40,48	95,73	57,45	11	28,16	47,67	265,58	62,86
Cobertura dos ativos não correntes (%)	11	203,47	254,83	453,34	468,11	11	186,90	266,70	306,20	413,60
Financiamento										
Peso do passivo remunerado (%)	11	0,00	0,00	5,21	6,52	11	0,00	2,53	5,49	5,75
Rendibilidade										
Rendibilidade dos capitais próprios (%)*	11	30,05	52,46	53,21	41,62	11	22,59	42,12	48,81	32,42
Efeito da atividade operacional	11	9,70	17,06	17,35	13,71	11	6,76	12,91	20,17	11,78
Efeito das atividades financeiras	11	2,78	3,10	6,21	3,11	11	1,48	2,49	5,62	2,77
Efeito da atividade não operacional	11	0,98	0,99	0,93	0,97	11	0,93	1,00	3,57	1,01
Efeito fiscal	11	1,00	1,00	0,97	0,98	11	1,00	1,00	0,96	0,98
Efeito da atividade extraordinária	11	1,00	1,00	1,00	1,00	11	1,00	1,00	1,00	0,99
Rendibilidade do ativo (%)	11	30,05	52,49	56,53	80,15	11	22,59	42,12	54,76	83,41
Rendibilidade das vendas (%)	11	12,89	17,49	40,99	36,58	11	4,02	17,40	38,74	34,87
VAB em percentagem da produção (%)	11	40,99	62,98	55,45	73,70	11	37,13	66,11	58,43	77,67
Risco										
Grau de alavancagem combinada**	11	0,24	1,62	1,35	1,15	11	1,04	1,60	1,62	1,47
Grau de alavancagem operacional	11	0,16	1,55	0,96	1,04	11	1,28	1,86	1,73	1,33
Grau de alavancagem financeira	11	0,95	1,06	1,03	1,07	11	0,96	1,07	-0,66	1,10
Grau de alavancagem da atividade não operacional	11	1,00	1,01	1,13	1,03	11	1,00	1,00	0,96	0,99
Grau de alavancagem da atividade fiscal	11	1,00	1,00	1,05	1,01	11	1,00	1,00	1,06	1,01
Grau de alavancagem da atividade extraordinária	11	1,00	1,00	1,00	1,00	11	1,00	1,00	1,00	1,00
Atividade										
Prazo médio de rotação dos inventários (nº dias)	11	55,09	90,31	204,09	171,30	11	15,66	44,22	81,64	112,71
Rotação do ativo (nº vezes)	11	0,47	0,66	0,62	0,70	11	0,52	0,69	0,68	0,89

Legenda: Nº Emp. – Número de Empresas; 1ºQ – 1º Quartil; 2ºQ – 2º Quartil; 3ºQ - 3º Quartil. Fonte: Banco Nacional de Angola

Nota: Rendibilidade dos capitais próprios (%)* , Grau de alavancagem combinada** e seus efeitos - Representam a média ponderada do agregado

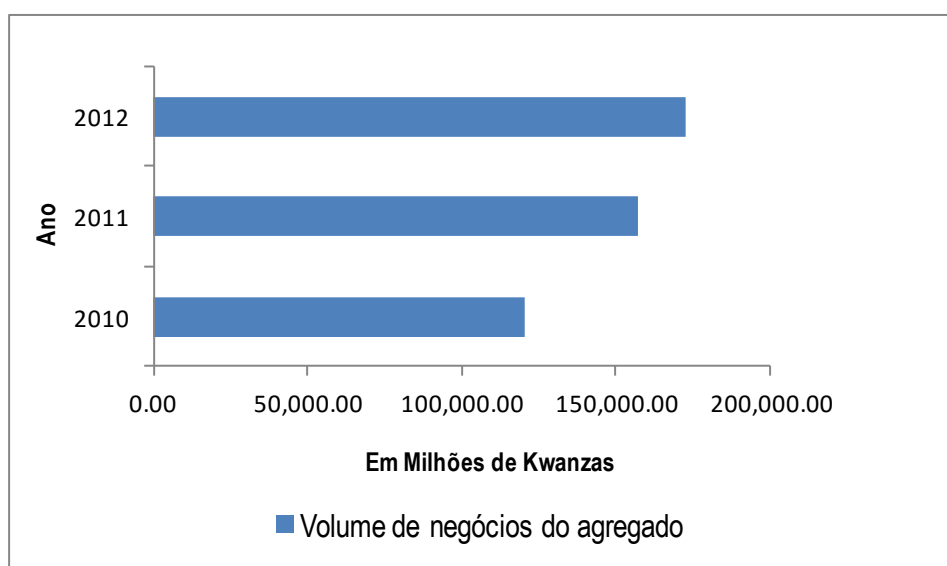
11.1. Atividade e Rendibilidade

Nesta secção serão analisados alguns indicadores que influenciam a atividade e rendibilidade das empresas, designadamente o volume de negócios, os custos da atividade operacional e a rendibilidade dos capitais próprios.

11.1.1. Volume de Negócios

O gráfico 3 permite analisar o comportamento do volume de negócios do agregado, o qual é constituído pelo conjunto das 11 empresas que integram a amostra em estudo.

Gráfico 3. Volume de Negócios do agregado das empresas em análise



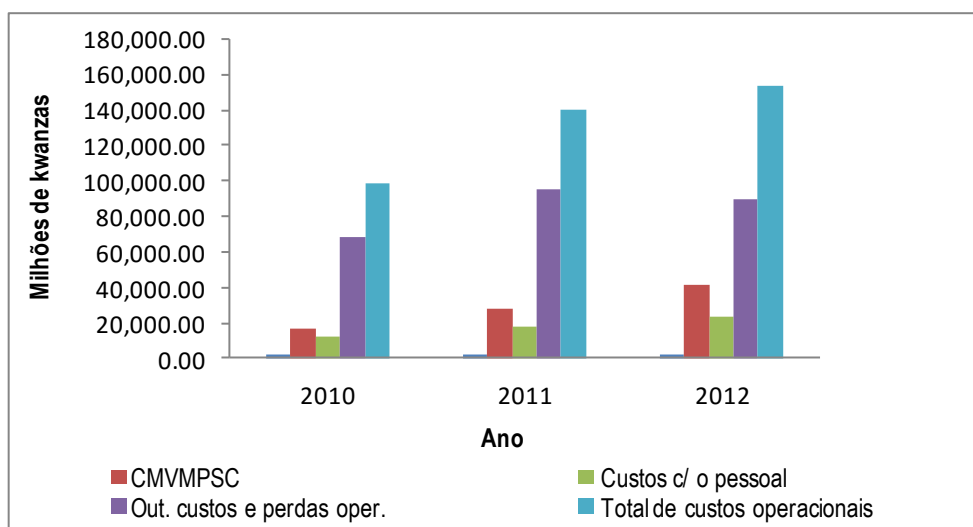
Fonte: BNA

A análise dos dados subjacentes ao Gráfico 3 permitiu verificar que, em 2012, o volume de negócios do agregado sofreu uma expansão de 9,84%, ao passar de 156.866,74 Milhões de Kwanzas em 2011 para 172.310,54 Milhões Kwanzas em 2012. Esta expansão, deveu-se ao facto de que 36,30% das empresas do agregado (4 empresas) sofreram um decréscimo no seu volume de negócios. De salientar que, estas 4 empresas representavam cerca de 49,90% do volume total de negócios do agregado em 2012.

11.1.2. Custos da Atividade Operacional

No Gráfico 4 estão representados os custos de atividade operacional para os três anos considerados no estudo (2010, 2011 e 2012).

Gráfico 4. Custos da atividade operacional



Fonte: BNA

O total de custos de atividade operacional passaram de 140.320,48 Milhões de Kwanzas em 2011 para 154.139,65 Milhões de Kwanzas, o que se traduz num aumento de 9,85%. Este deveu-se principalmente ao crescimento do custo das mercadorias vendidas e das matérias-primas e subsidiárias consumidas (CMVMPSC), na ordem de 48,35%, e dos custos com o pessoal de 33,99%, apesar do decréscimo verificado nos outros custos e perdas operacionais (Out. custos e perdas oper.) ter sido cerca de 5,89%.

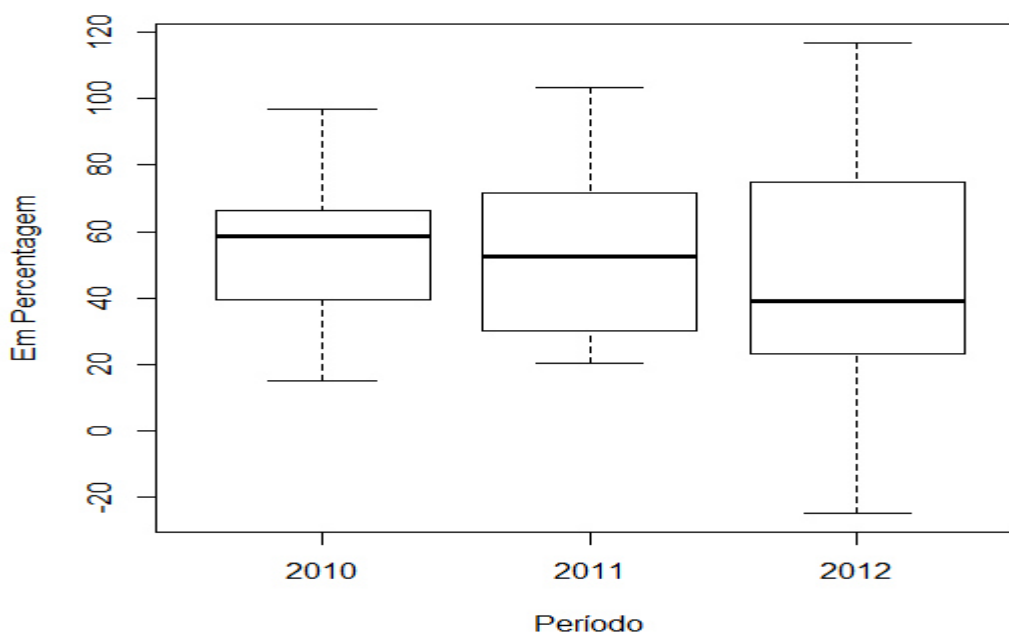
Em relação à estrutura de custos, verificou-se que em 2012 os CMVMPSC representaram 26,62% dos custos operacionais, enquanto que os custos com pessoal e os outros custos representaram 15,43% e 57,95%, respetivamente. De salientar, que se verifica um desequilíbrio, significativo na estrutura de custos operacionais deste agregado, pelo que se deve aqui equacionar uma reestruturação de custos, principalmente inerentes ao pessoal, que se situam abaixo dos 20,00%.

De um modo geral, os custos de atividade operacional das empresas do setor de construção têm vindo a aumentar. Este aumento pode estar associado ao aumento da concessão de obras públicas e privadas, que implica a contratação de mão-de-obra especializada e importação de materiais de construção, sujeitos ao aumento dos custos com transporte e direitos aduaneiros em Angola.

11.1.3. Rendibilidade dos Capitais Próprios

No Gráfico 5 estão representadas as caixas de bigode (*boxplot's*) referentes aos três anos em análise. Este tipo de representação gráfica realça algumas características dos dados, tais como as medidas de localização e dispersão, nomeadamente a mediana e os quartis, respetivamente.

Gráfico 5. Rendibilidade dos Capitais Próprios



Fonte: BNA

Os diagramas de caixa permitiram verificar que a mediana associada à rendibilidade dos capitais próprios do agregado sofreu uma diminuição de 2010 para 2011 e de 2011 para 2012. Pode observar-se que as distribuições das medidas de localização nos anos 2010 e 2011 apresentam algumas semelhanças, isto é, as medianas aproximam-se do terceiro quartil, o que indica uma distribuição assimétrica negativa. Contrariamente, em 2012, a distribuição é assimétrica positiva, visto que a mediana surge próxima do primeiro quartil. Tal resulta da discrepância na rendibilidade dos capitais próprios do agregado em 2012, uma vez que os valores dos extremos inferiores e superior surgem muito afastados. Observou-se ainda que o valor da mediana e o valor do primeiro quartil da rendibilidade dos capitais próprios das empresas do agregado registaram valores tendencialmente inferiores ao longo dos três momentos de análise.

No Quadro 14, observa-se que a rentabilidade média do agregado dos capitais próprios passou de 41,62% em 2011 para 32,42% em 2012, significando uma redução de 9,20 pontos percentuais. Com exceção do efeito da atividade não operacional, todos os outros efeitos sofreram reduções comparativamente a 2011, enquanto a liderança de contribuição é mantida pelo efeito da atividade operacional com 11,79%, seguido do efeito da atividade financeira com 2,77%.

11.2. Situação Financeira

11.2.1. Estrutura Financeira

O Quadro 15 apresenta a distribuição por quartis e a média agregada do rácio de autonomia financeira.

Quadro 15. Autonomia Financeira: Distribuição por quartis e média agregada

Ano	Q1	Q2	Q3	Média do Agregado
2010	20,5%	28,6%	38,6%	50,3%
2011	21,4%	28,8%	34,2%	36,5%
2012	21,9%	31,7%	38,8%	38,6%

Fonte: BNA

O rácio de autonomia financeira mede a solvabilidade da empresa através da determinação da proporção dos ativos que são financiados com capital próprio. Este rácio é utilizado pelos bancos comerciais para avaliar o equilíbrio financeiro de uma empresa, quando a mesma recorre ao sistema bancário para obter um financiamento de curto prazo para necessidades de tesouraria. Um rácio baixo indica uma maior vulnerabilidade financeira.

Pode concluir-se, pelo Quadro 15, que a média do rácio de autonomia financeira do agregado sofreu um aumento entre 2011 e 2012 ao passar de 36,5% para 38,6%, respetivamente, representando uma variação de 2,10 pontos percentuais. De um modo geral, existe um crescimento, ao longo dos anos, do rácio de autonomia financeira (as medidas de localização aumentaram), isto é, a solidez financeira do agregado aumentou bem como a capacidade para solver os compromissos não correntes.

Este comportamento pode estar associado ao facto de haver um maior recurso aos financiamentos da casa mãe ou empresas do grupo, por se tratarem de empresas de investimento direto estrangeiro em Angola, uma vez que o rácio do peso do passivo remunerado passou de 6,52%, em 2011, para 5,75%, em 2012, ou seja, houve uma redução de 0,80 pontos percentuais.

No entanto, a metodologia recomenda que este rácio, à semelhança de outros, não seja analisado isoladamente. Assim, propõe-se a sua associação a indicadores como o prazo médio de recebimentos, prazo médio de pagamentos, e prazo médio de rotação de inventário, entre outros indicadores.

Com base nos indicadores económico-financeiros dos rácios de atividade, foi usado no caso prático apenas o prazo médio de rotação de inventário, que corresponde ao número de dias que a empresa demora a consumir ou vender o inventário. Concluiu-se que em 2012 as empresas demoravam cerca de 113 dias a converter o seu *stock* em vendas, um período mais curto comparativamente a 2011, cujo prazo era de 171 dias. Verificou-se assim uma melhoria de cerca de 60 dias, o que representa mais de 30% de redução (Quadro 14).

11.2.2. Custos Financeiros e Solvabilidade

Tratando-se de empresas de investimento direto, as transações de financiamentos entre a casa matriz e as suas filiais são tratadas como investimento direto, o que se pode considerar como sendo uma das razões pelas quais os indicadores financeiros não refletem custos financeiros.

Com vista a avaliar a capacidade da empresa cumprir com as obrigações de curto prazo, se analisou a liquidez geral e a liquidez reduzida do conjunto das empresas que constituem o agregado.

Com base nos indicadores económico-financeiro (Quadro 14), verifica-se que a liquidez geral do agregado atingiu os 153,0% em 2012, comparativamente aos 145,8% de 2011.

O rácio da liquidez reduzida, também conhecido como *acid test*, mede a liquidez de forma mais restritiva, ou seja, exclui os inventários. Este indicador favorece o setor da construção em que a liquidez dos inventários é menor e a conversão dos artigos em *stock* em dinheiro poderá ser mais difícil. À semelhança da liquidez geral, a liquidez reduzida sofreu um aumento ao passar de 144,6% em 2011 para 152,5% em 2012, verificando-se assim a melhoria da capacidade das empresas do agregado pagarem as suas dívidas de curto prazo apenas com as disponibilidades e com o valor que os clientes devem à empresa (Quadro 14).

Por outro lado, promoveu-se a análise do rácio de solvabilidade geral, que permite avaliar a estrutura de financiamento da empresa, colocando em evidência o peso dos capitais investidos pelos sócios ou acionistas no total dos capitais alheios (provenientes de entidades externas).

Este rácio situou-se em 57,45% em 2011 e passou para 62,86% em 2012. O que demonstra que o agregado está solvente do ponto de vista económico, pois apresenta um capital próprio que garante a liquidação do seu passivo e possui expectativas de resultados que garantam a sua sobrevivência futura. O agregado está em condições de fazer face às suas responsabilidades correntes.

11.3. Análise de Risco Económico-financeiro

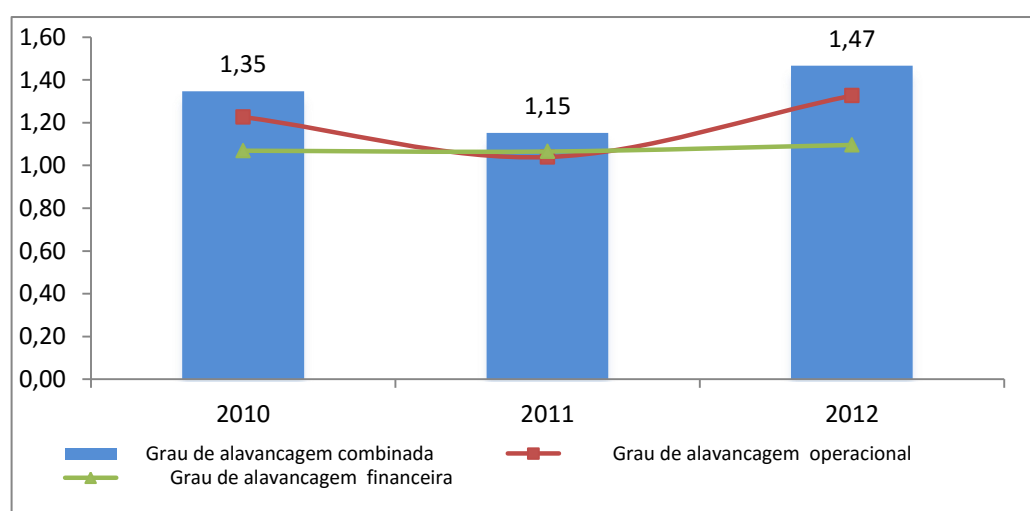
Para avaliar o risco económico e financeiro do agregado foi analisado o grau de alavancagem combinada, que passou de 1,15% em 2011 para 1,47% em 2012, ou seja, verificou-se um aumento de 0,31 pontos percentuais (Gráfico 6).

Este resultado foi influenciado maioritariamente pelos graus de alavancagem operacional e financeira que sofreram aumentos de 0,29 pontos percentuais e 0,03 pontos percentuais em 2012 comparativamente a 2011, respetivamente.

Os graus de alavancagem da atividade fiscal, da atividade não operacional e da atividade extraordinária, em 2012, mantiveram um comportamento estável comparativamente a 2011, evolução esta que pode ser analisada através do Quadro 14.

Ao aumento do risco de negócio, podem estar associadas fatores como a redução das encomendas dos clientes, a maior exigência dos fornecedores quanto às condições de pagamento, eventuais dificuldades em determinados abastecimentos e a saída de alguns colaboradores.

Gráfico 6. Grau de alavancagem combinada



Conclusões e Recomendações

A análise de conglomerados do conjunto do setor não petrolífero em Angola (excluindo o setor financeiro) permitiu concluir que as empresas ativas não estão distribuídas de forma equitativa e uniforme pelos setores de atividade económica, pois mais de 50% das empresas ativas encontram-se inseridas no setor do comércio a grosso e a retalho. Esta situação deve-se ao fraco desenvolvimento de setores como a agricultura, as pescas, a indústria transformadora, e o setor da eletricidade, gás e água, o que expressa a dependência da economia angolana face às importações, com vista à satisfação da procura doméstica e, simultaneamente, para colmatar o fraco desenvolvimento da produção interna.

A identificação de *clusters* constitui o conhecimento de base para a implementação das medidas de política que visam o crescimento harmonioso e sustentável da economia nacional. À luz do Programa Nacional de Desenvolvimento 2013-2017, o Executivo definiu *clusters* prioritários para a diversificação da economia angolana, nomeadamente, o cluster da energia e água, da alimentação e agro-indústria, da habitação e o dos transportes e logística (República de Angola, 2012).

A análise das correlações revelou que a generalidade dos setores não está a beneficiar do desenvolvimento da agricultura (uma vez que apresentam com este setor correlações abaixo de 0,3) o que aponta para a necessidade de um maior investimento e incentivo a este setor.

Até ao conflito armado pós-independência, Angola era um país auto-suficiente na generalidade dos alimentos (à exceção do trigo) e foi, inclusive, um grande exportador de café e de milho, além de banana, a cana-de-açúcar, óleo de palma, coco e amendoim, feijão, mandioca. Entre as culturas comerciais destacaram-se o algodão e a borracha. Angola foi, inclusive, um dos maiores exportadores mundiais de café e de outros produtos agrícolas, como o algodão, sisal, milho, mandioca e banana. Desde o pós-independência, o país diminuiu a sua capacidade produtiva, em relação aos anos anteriores. Hoje a agricultura em Angola caracteriza-se por produções agrícolas muito aquém das suas potencialidades, obrigando o país a utilizar consideráveis recursos financeiros para a importação de alimentos.

Esta situação é possível de inverter no futuro uma vez que o país possui recursos naturais, nomeadamente terras aráveis e recursos hídricos, que permitiriam um desenvolvimento sustentável do setor agropecuário (Pacheco, Carvalho & Henriques, 2013). O desenvolvimento da agricultura permitiria, por sua vez, potenciar o desenvolvimento da indústria alimentar e têxtil. De acordo com a

Organização das Nações Unidas, Angola tem 46,3% das terras agrícolas, 575.900 km², mas apenas 5,7% é explorado. A existência de solos férteis, ricos em biodiversidade, boas condições climáticas, abundância de recursos hídricos e recursos naturais, combinada com uma população rural dinâmica cuja atividade está diretamente relacionado com a agricultura, a pecuária e as pesca, são fatores favoráveis para o desenvolvimento económico e social do país (FAO, 2012).

O investimento na agricultura representaria, numa primeira perspetiva, a satisfação da procura interna e a substituição paulatina do consumo das importações, e numa fase posterior a diversificação da estrutura das exportações, uma vez que o país seria mais autónomo e progressivamente teria mais excedentes para a exportação.

Sabe-se que uma economia de mercado alicerçada no tecido empresarial nacional é determinante no processo de diversificação económica, enquanto promotora da criação de emprego e da inovação das cadeias produtivas. Para tal, têm sido implementadas medidas de estímulo à economia, como a captação de investimento (nacional e estrangeiro), a concessão de financiamento às empresas, sobretudo nos setores prioritários, o programa Angola Investe, criado em 2012, e que tem por objetivo a concessão de financiamento às micro, pequenas e médias empresas, bem como a atribuição de incentivos fiscais e de facilidades no acesso ao crédito. Estes programas visam fortalecer o tecido empresarial nacional, aumentar a produção doméstica para reduzir a dependência nacional das importações e diversificar a estrutura das exportações e as fontes de receitas fiscais e de divisas para fazer face às oscilações do preço de petróleo no mercado internacional, e finalmente promover a criação de emprego e a integração profissional dos quadros angolanos.

O conhecimento integral e profundo sobre as similaridades e os pontos de convergência dos diferentes setores permite compreender as dinâmicas económicas do país, bem como identificar sinergias e complementaridades entre setores que podem ser potenciadas de forma a alcançar um desenvolvimento conjunto.

Uma vez que a agricultura apresentou uma correlação moderada e positiva com o setor das pescas, concluiu-se que o incentivo e a dinamização da agricultura beneficiariam também as pescas e vice-versa. O setor das pescas, por sua vez, apresentou uma correlação positiva com o conjunto dos setores da eletricidade, gás e água, da construção, do comércio, grosso e retalho, dos transportes e armazenagem e da atividade imobiliária, os quais apresentaram também correlações fortes e positivas entre si. Este resultado indica que uma evolução positiva ao nível das pescas beneficiaria os restantes

setores de forma proporcional. Em Angola, o relançamento deste setor consta do Plano de Ordenamento das Pescas e Aquicultura (República de Angola, 2012) que visa potenciar a extensa linha costeira angolana com cerca de 1.650 km, através da concessão de financiamento e formação das pessoas para renovação da frota e das metodologias de atuação.

O desenvolvimento da indústria transformadora irá beneficiar os setores que se encontrem a jusante e a montante do setor industrial, como a agricultura (em termos de matérias primas agrícolas e máquinas, equipamentos e outros *inputs* para a atividade agrícola), as pescas (em termos de equipamentos para a atividade pesqueira) e os transportes e comunicações (por meio do transporte de máquinas, equipamentos e matérias primas para a indústria e os meios de transportes). O setor da energia e água representam a disponibilidade de *inputs* para a atividade industrial; os setores da construção e do imobiliário permitiram a definição e orientação de pólos de desenvolvimento industrial.

O conjunto dos setores da indústria transformadora, hotelaria, alojamento e restauração, da educação, da saúde e ação social, e outras atividades apresentaram uma correlação negativa com o setor das pescas e uma correlação também negativa entre si. Este resultado permite concluir que a evolução destes setores tem ocorrido de forma inversamente proporcional.

Na análise de correlações destaca-se ainda a relação forte e positiva entre o setor da construção e os setores do imobiliário, do comércio a grosso e retalho, dos transportes e armazenagem e da eletricidade, gás e água. A construção de novas urbanizações, a requalificação do parque habitacional e a construção de projetos de habitação social integram o conjunto das medidas do Executivo para promover o acesso à habitação própria e, ao mesmo tempo, estimulam o setor da construção e do mercado imobiliário. No mesmo sentido, tem sido reforçado o investimento na construção de obras públicas (como escolas, hospitais e outras) e vias de comunicação com o intuito de criar as condições de base para a circulação de bens e pessoas em todo o território nacional. Desta forma são potenciados os setores dos transportes e armazenagem e do comércio a grosso e retalho. A partir de um maior investimento em infra-estruturas energéticas e de telecomunicações é alcançado o desenvolvimento da indústria, a modernização dos centros industriais e de logística, e o apoio ao comércio a grosso e retalho. No entanto, esta é uma dimensão do desenvolvimento e da reconstrução nacional ainda em curso e para a qual se tem procurado atrair investimento externo e canalizar o trabalho de empresas estrangeiras do setor da construção, como por exemplo Portugal, China e Brasil. Angola é hoje, um país para o qual o apoio financeiro e a assistência técnica se apresentam como alavancas para a reconstrução das infra-estruturas necessárias à competitividade da sua economia.

A formação dos conglomerados através do critério *average linkage* e a medida de dissimilaridade baseada na correlação permitiu obter três *clusters* de setores coincidentes com a tendência observada através da matriz de correlações. Assim, obteve-se um primeiro *cluster* que integrava apenas o setor da agricultura, um segundo *cluster* que integrava o setor das pescas e os demais que se correlacionavam positivamente entre si, e um terceiro *cluster* que integrava os setores com coeficientes de correlação negativos. Este último *cluster* manteve a mesma composição através do critério *single linkage* (e medida baseada na correlação) a par de um segundo setor que integrava todos os restantes setores, incluindo a agricultura. Os resultados de conglomerados pelo critério *average linkage* e a dissimilaridade baseada na correlação confirmou também as suposições levantadas na sequência da análise à evolução do número de empresas por setores de atividade económica, sendo que o segundo conglomerado (A_2) integrou os setores com evolução crescente e o terceiro conglomerado (A_3) agrupou os setores com evolução decrescente. A tendência semelhante observada na análise descritiva entre os setores da agricultura, pescas e eletricidade, gás e água foi evidente na formação do segundo conglomerado pelo critério *simple linkage* (B_2) e do primeiro conglomerado critério *average linkage* e a dissimilaridade baseada no periodograma simples (C_1).

Quando estudados do ponto de vista da sazonalidade (através da medida de dissimilaridades do periodograma simples) o critério *average linkage* permitiu a formação de quatro grupos de setores, enquanto o critério *complete linkage* permitiu a constituição de cinco grupos de setores. Os setores da construção e da atividade imobiliária constituíram, cada um, um *cluster* independente. Os setores da indústria transformadora e do comércio, grosso e retalho integraram um único *cluster*, enquanto que os restantes setores de atividade económica formaram um outro *cluster* obtido pelo critério *average linkage*. O critério *complete linkage* permitiu a constituição de cinco conglomerados, de constituição muito semelhante à observada pelo critério ligação média, com a diferença de que o setor da hotelaria, do alojamento e restauração integrou um único setor. Conclui-se assim que em termos de sazonalidade existem características que tornam os setores muito dissemelhantes entre si.

Os grupos obtidos pelos diferentes critérios (com base na correlação e com base no periodograma simples) registaram dimensões muito diversas, pelo que se considera que os resultados obtidos não correspondem a resultados de conglomerados equilibrados. Contudo, a análise de conglomerados permitiu identificar grupos de setores com características semelhantes entre si e dissemelhantes em relação aos outros grupos, compreender como os vários setores se correlacionam e como a evolução de um setor afeta os restantes.

Uma vez que em 2012 o setor da construção integrou a terceira posição no PIB não petrolífero em Angola, de acordo com o Contas Nacionais 2007-2012 (INE Angola, 2014) e registou um crescimento de 21,8% (face ao valor de 12% de 2011) foi escolhido este setor para proceder a uma análise micro da evolução do agregado do setor. Para tal, foi constituída uma amostra de 11 empresas portuguesas do setor da construção a operar em Angola (de 2010 a 2012) que permitiu a aplicação da metodologia do BdP para a produção das estatísticas das SNF.

Face ao problema de não resposta, foi necessário proceder à imputação dos dados omissos, através do método de imputação pela média e do método Cold-Deck. Depois da compilação dos dados contabilísticos e da imputação dos dados omissos, foi obtido o quadro de setor que possibilitou a análise dos indicadores económicos e financeiros da construção. A compilação e a produção das estatísticas aplicadas ao setor da construção permitiram a obtenção do quadro de setor e a análise dos indicadores económicos e financeiros do mesmo. Os resultados revelaram o crescimento do volume de negócios, em concordância com os resultados atingidos para o setor como um todo em 2012 e segundo a informação publicada no IAHE 2009-2012 (INE Angola, 2014).

A compilação e produção das estatísticas das SNF permite, assim, obter indicadores de avaliação da evolução da atividade, rendibilidade e património das empresas, informação fundamental na formulação das políticas públicas setoriais e macroeconómicas, nomeadamente em termos dos programas de investimento que Angola tem em curso.

O estágio curricular realizado na Central de Balanços do BdP sobre a compilação das estatísticas das SNF possibilitou o conhecimento da experiência e da metodologia do BdP na compilação de dados contabilísticos e na produção das estatísticas das SNF. Por outro lado, o facto de Portugal integrar o conjunto dos 11 países membros do BACH significou a oportunidade de conhecer a metodologia de aplicação das normas contabilísticas internacionais da zona euro.

Para Angola, este conhecimento constitui uma vantagem, no sentido em que poderá servir de alavancagem para proceder à implementação, a curto e médio prazo, de uma Central de Balanços com base na compilação de informação contabilística anual. A seu tempo permitirá, ainda, alargar a atual cobertura do IAHE, em vez de considerar uma amostra de empresas nacionais, proceder à recolha de informação estatística do universo das SNF.

Numa perspetiva de longo prazo, a recolha e compilação de dados trimestrais sobre a atividade das empresas nacionais, a par do conhecimento macro dos setores de atividade económica, representa a possibilidade de obter estatísticas micro e macro do agregado das empresas que atuam em Angola. A produção de estatísticas trimestrais permitirá a obtenção de um conhecimento atualizado da estrutura e

evolução da atividade empresarial, fundamental para a implementação de políticas públicas orientadas para uma maior competitividade da economia e monitorização da atividade económica e financeira. Espera-se assim que o estudo do caso prático desenvolvido no BdP constitua um incentivo de mudança (ainda que modesto) no sentido de serem equacionadas novas práticas da compilação e publicação das estatísticas das SNF em Angola.

Através do estudo da metodologia em uso no BdP foi possível conhecer as vantagens e mais-valias de proceder a uma recolha da informação contabilística das empresas em suporte informático, uma vez que são evitadas perdas de informação e torna possível uma gestão mais eficiente dos dados. Por isso, propõem-se que a informação recolhida pelo IAHE e pelo Ministério das Finanças seja armazenada em suporte eletrónico, de forma a salvaguardar os dados e a qualidade de informação das estatísticas das SNF. Para garantir a qualidade da informação, poderá proceder-se ao controlo de qualidade, com recurso as bases de dados geridas pelo BNA, Ministério da Justiça (registo comercial), Agência Geral Tributária do Ministério das Finanças, Ficheiro de Unidades Empresariais do INE Angola, outras bases de dados disponíveis e Relatório e Contas publicados pelas SNF, consideradas como potenciais fontes internas ou externas. Por outro lado, considera-se pertinente a utilização de uma ferramenta compatível com o volume de informação a tratar e a sua dimensão confidencial.

Outras recomendações que decorrem do caso prático prendem-se com a recolha da informação do balanço e dos fluxos de caixa das SNF de forma a permitir, numa fase transitória, a produção de estatísticas das empresas e a análise económica, financeira e patrimonial.

Recomenda-se também que a informação sobre os segmentos de atividade económica e as classes de dimensão (que consta atualmente do FUE) seja disponibilizada de forma a permitir uma análise mais segmentada e, portanto, mais profunda em termos da estrutura e evolução do setor. Seria igualmente oportuno estudar a localização geográfica das empresas, de forma a avaliar a concentração e a distribuição por Província e Distrito, por exemplo, e nesta base, poder equacionar políticas de investimento e de redistribuição da atividade das empresas pelas regiões menos potenciadas.

Por último, a produção e divulgação das estatísticas das SNF poderá ser realizada de forma atualizada e oportuna para possibilitar uma monitorização permanente da atividade económica e financeira do tecido empresarial angolano, e assim reduzir o desfasamento temporal das últimas publicações.

Para futuras investigações recomenda-se a análise económica e financeira com base em dados do conjunto de todas as empresas ativas num determinado setor, ou de um conjunto de setores considerados cruciais para o desenvolvimento da economia angolana. Considera-se igualmente pertinente proceder a uma análise de clusters para um conjunto de dados de um maior período temporal, de forma a formular previsões sobre a evolução dos setores.

Na continuidade da análise de conglomerados perspetiva-se um estudo mais alargado que considere outras variáveis como o volume de negócios e o número de pessoas ao serviço em cada setor de atividade com vista a sua caracterização com base nas semelhanças e dissemelhanças, para a implementação de medidas de política. Por outro lado a aplicação de outras técnicas de análise exploratória e de reconhecimento de padrões poderão para permitir a formulação de hipóteses e a previsão do comportamento dos setores.

Não obstante terem sido alcançados os objetivos do estudo, surgiram ao longo da investigação algumas limitações e constrangimentos que obrigaram a reequacionar a metodologia de estudo. A principal limitação do estudo do caso prático residiu na impossibilidade de acesso à informação micro das empresas não financeiras do setor da construção. Esta informação é publicada de forma agregada pelo INE de Angola, e por todas as instituições estatísticas que compilam e divulgam estatísticas. Esta limitação está associada ao princípio do Segredo Estatístico e a Lei de Bases do Sistema Estatístico Nacional (Lei nº 6/89, de 15 de Abril), o que constituiu uma limitação para o estudo inicialmente planeado, o qual visava a análise de um ou dois setores da economia Angolana. Tal limitação impossibilitou o estudo do setor da construção na íntegra e obrigou à compilação de dados de forma manual de um conjunto de onze empresas, recorrendo a uma fonte do BNA. Por outro lado, apenas foi possível o acesso a relatórios consolidados do agregado. Assim, o estudo não teve a abrangência desejada que seria a de cobrir na totalidade, pelo menos, um dos setores da atividade económica de Angola.

Contudo, este estudo acaba por constituir um ponto de partida na análise das estatísticas das SNF, com o propósito mais vasto de demonstrar a importância e as potencialidades de dispor de um conhecimento alargado sobre todos os setores de atividade económica em Angola.

Referências

- [1] Banco de Portugal (2009). *Estatística das Sociedades Não Financeiras da Central de Balanços. Documento Metodológico*. Departamento de Estatística. Área da Central de Balanços.
- [2] Banco de Portugal (2013). *Boletim Estatístico 5 | 2013*. Banco de Portugal. Eurosistema.
- [3] Banco de Portugal (2013). *Estatísticas das empresas não financeiras da central de balanços – notas metodológicas. Suplemento 2/2013 ao Boletim Estatístico*. Banco de Portugal Eurosistema.
- [4] Banco de Portugal (2014). *Quadros do setor e quadros da empresa e dos setores. Notas Metodológicas Série longa 1995-2013*. Estudos da Central de Balanço. Eurosistema.
- [5] Banco Mundial (2013). *Angola Economic Update*, 1. Junho. Banco Mundial.
- [6] Banco Nacional de Angola (s/d). *Relatório e Contas 2012*. Luanda: Banco Nacional de Angola.
- [7] Belcher, R. (2003). Application of the *Hidiroglou-Berthelot* method of outlier detection for periodic business surveys. *Proceedings of the Survey Methods Section*. SSC Annual Meeting, June.
- [8] Busaab, W.; Miazaki, E.S. & Andrade, D. F. (1990). *Introdução à análise de agrupamentos*. São Paulo: Associação Brasileira de Estatística.
- [9] Carr, D. B.; Young, C. J.; Aster, R. C. & Zhang, X. (1999). *Cluster Analysis for CTBT Seismic Event Monitoring*. Conference: 21st Seismic Research Symposium: Technologies for Monitoring the CTBT, Las Vegas.
- [10] Denton, F. T. (1971). Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization. *Journal of the American Statistical Association* 66, 333 (3), 99-102.
- [11] European Banking Authority (2012). *Assessment of SME proposal for CRD IV/CRR*. London: EBA.
- [12] Eurostat (s/d). *Recommended Practices for Editing and Imputation in Cross-Sectional Business Survey*.
- [13] Everitt, B.S.; Landau, S.; & Leese, M. (2001). *Cluster Analysis*. 4th Edition. London: Arnold.
- [14] Food and Agriculture Organization (2012). *Angola County Programming Framework 2013-2017*. Luanda: FAO.
- [15] Galton, Francis. (1869). *Hereditary Genius*.

- [16] Instituto Nacional de Estatística de Angola (2014). *Contas Nacionais 2007-2012*. Outubro de 2014. Luanda: INE Angola.
- [17] Instituto Nacional de Estatística de Angola (2015). *Estatísticas do Ficheiro de Unidades Empresariais. FUE 2010-2013*. Luanda: INE Angola.
- [18] Instituto Nacional de Estatística Angola (2015). *Inquérito Anual Harmonizado às Empresas 2012*. Luanda: INE Angola.
- [19] Instituto Nacional de Estatística de Portugal (2012). *Documento Metodológico – Inquérito Trimestral às Empresas Não Financeiras, versão 2.1*. Lisboa: INE Portugal.
- [20] International Monetary Fund (FMI) (2013). *World Economic Outlook, 2013*. International Monetary Fund.
- [21] Little, R. (1992). Regression with Missing Xs – A Review. *Journal of the American Statistical Association* 87(420): 227-37
- [22] Little, R. J. A. (1988). Missing-data adjustments in large surveys. *Journal of Business and Economic Statistics*.
- [23] Marôco, J. (2011). *Análise Estatística com o SPSS Statistics (5ª ed.)*. Lisboa, Edições Report Number.
- [24] McQuitty, L. L. (1964). Capabilities and Improvements of Linkage Analysis as a Clustering Method. *Educational and Psychological Measurement* 24(3): 441-456.
- [25] Mukaka, M. M. (2012). A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal* 24(3): 69-71.
- [26] Nunes, L. N.; Klück, M. M.; & Fachel, J. M. G. (2010). Comparação de métodos de imputação única e múltipla usando como exemplo um modelo de risco para mortalidade cirúrgica. *Revista Brasileira de Epidemiologia* 13(4): 596-606.
- [27] Pacheco, F.; Carvalho, L. S.; & Henriques, D. S. (2013). Contribuição para o debate da sustentabilidade da agricultura angolana. *Economia, Sociologia, Ambiente e Desenvolvimento Rural - Atas do 2.º Encontro Luso-Angolano na Universidade Metodista de Angola*. Luanda, 6 a 8 de Outubro 2011: 311-343.
- [28] Pinto, R. R. (2009). *Introdução à Análise de Dados Com Recurso ao SPSS*.
- [29] Reis, E. (2001). *Estatística Multivariada Aplicada*, 2ª ed. Lisboa: Edições Sílabo.

- [30] República de Angola (2012). Programa Nacional de Desenvolvimento 2013-2017. Luanda: Ministério do Planeamento e do Desenvolvimento Territorial.
- [31] Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.
- [32] Schultz & Schultz. (1992). Teoria da regressão.
- [33] Sokal, R. R. & Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *Taxon*, 11: 33-40.
- [34] Sokal, R. R. & Sneath, P. H. (1963). *Numerical Taxonomy. The principles and the practice of numerical classification*. São Francisco: W. H. Freeman and Company.
- [35] Suplemento 5/2005 ao Boletim Estatístico. Lisboa: INE Portugal.
- [36] Tukey, John Wilder . (1977). *Exploratory Data Analysis*.
- [37] Webb, A. W. (2002). *Statistical Pattern Recognition*, 2nd Edition. Malvern: John Wiley & Sons, Ltd.
- [38] Young, W. , Weckman, G. and Holland, W.(2011). A survey of methodologies for the treatment of missing values within datasets: limitations and benefits, *Theoretical Issues in Ergonomics Science*.
- [39] Zhang, P. (2003). Multiple imputation: theory and method. *International Statistical Review* 71(3): 581-92.

Legislação Angolana:

- [40] Decreto nº82/01, de 16 de novembro. Plano Geral de Contabilidade de Angola.
- [41] Lei n.º 15/96, de 27 de setembro. Lei de Bases do Sistema Nacional de Estatística.
- [42] Lei n.º 16/10, de 15 de julho. Lei do Banco Nacional de Angola.
- [43] Lei n.º 30/11, de 13 de setembro. Lei das Micro, Pequenas e Médias Empresas.

Legislação Portuguesa:

- [44] Decreto-Lei nº 8/2007, de 17 de janeiro. Cria a Informação Empresarial Simplificada (IES).
- [45] Decreto-Lei nº 381/2007, de 14 de novembro, que substituiu a CAE-Rev.2.1, a partir de 1 de janeiro de 2008, pela Classificação Portuguesa de Atividades Económicas, Revisão 3 (CAE-Rev.3).

- [46] Decreto-Lei nº 136/2012, de 2 de junho de 2012. Lei Orgânica do Instituto Nacional de Estatística.
- [47] Lei nº 5/98, de 31 de janeiro. Lei Orgânica do Banco de Portugal.
- [48] Lei nº 22/2008, de 13 de maio de 2008. Lei do Sistema Estatístico Nacional.
- [49] Recomendação da Comissão Europeia, de 6 de maio de 2003, relativa à definição de micro, pequenas e médias empresas.

ANEXOS

Quadro 16. Caracterização do agregado do Quadro do Setor

QUADRO DO SETOR			
AGREGADO (CAE/Dimensão)			
Ano:	2014		
Setor de Atividade Económica (CAE Rev.3)	F- Construção		
Classe de Dimensão:	Grandes empresas		
A.CARATERIZAÇÃO DO AGREGADO (I)			
		2013	2014
Número de empresas incluídas no agregado			
REPRESENTATIVIDADE (Intervalos em percentagem)		2013	2014
Em Número de empresas		95 % - 100%	90 % - 95%
Em Número de pessoas ao serviço		95 % - 100%	95 % - 100%
Em Volume de negócios		95 % - 100%	95 % - 100%
MOVIMENTOS NO AGREGADO DO ANO MAIS RECENTE	Entrada de empresas	Saída de empresas	Empresas com acontecimentos marcantes
Em número de empresas	6,00	15	2

Fonte: Banco de Portugal

Quadro 17. Rácios Setoriais Europeus (a partir de 2006) - Base de Dados BACH para França e Portugal

QUADROS DO SETOR								
Ano:	2014							
Setor de Atividade Económica (CAE Ver.3)	F- Construção							
Classe de Dimensão:	Grandes empresas							
F. RÁCIOS ECONOMICO FINANCEIROS								
	2013							
	França				Portugal			
	Media	Quartil 1	Mediana	Quartil 3	Media	Quartil 1	Mediana	Quartil 3
RÁCIOS SETORIAIS EUROPEUS (DADOS A PARTIR DE 2006)								
Rendibilidade e atividade								
Volume de negócios / Total do ativo (%)	112,46	145,36	190,55	241,93	31,10	0,00	41,55	115,99
Gastos com o pessoal / VAB (%)	86,35	76,99	86,91	94,77	75,99	58,66	83,31	98,66
Resultado antes de impostos / Capital próprio (%)	20,20	4,71	15,70	32,82	-1,53	-2,95	1,43	12,47
Clientes / Volume de negócios (%)	27,60	14,36	21,48	29,36	47,39	1,42	23,82	58,56
VAB / Volume de negócios (%)	29,52	28,49	37,53	45,75	28,36	17,11	35,58	54,90
Resultado antes de impostos / Volume de negócios (%)	4,43	0,60	2,63	5,88	-1,08	-14,28	1,09	5,56
Estrutura das origens e decomposição dos resultados								
Capital próprio / Total do ativo (%)	24,64	19,43	34,21	49,45	22,00	0,52	24,41	56,35
Fornecedores / Total do ativo (%)	21,31	13,75	21,56	31,57	10,41	0,00	5,30	22,33
Total de rendimentos / Volume de negócios (%)	105,54	100,06	100,69	102,22	110,20	100,00	100,07	103,60
Total de gastos / Volume de negócios (%)	102,16	95,82	98,58	101,37	112,36	93,85	100,35	121,24
Estrutura do ativo								
Investimentos financeiros / Total do ativo (%)	8,10	0,10	0,47	1,33	14,39	0,00	0,00	0,00
Clientes / Total do ativo (%)	31,04	26,45	41,31	56,13	14,74	0,00	12,11	44,94
Caixa, dep bancários e instr. financeiros / Total do ativo (%)	15,23	4,06	15,21	32,53	5,85	0,53	4,52	20,88

Figura 9. Dissimilaridade baseada na autocorrelação parcial (método da média e método do centróide)

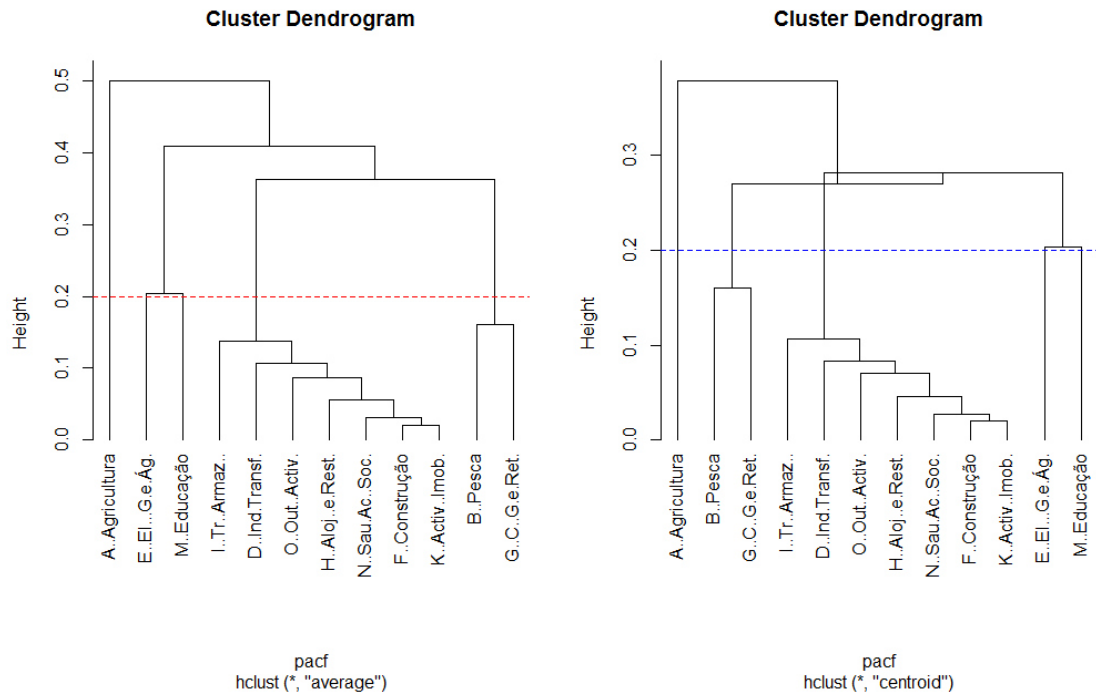


Figura 10. Dissimilaridade baseada na autocorrelação simples (método da média e método da mediana)

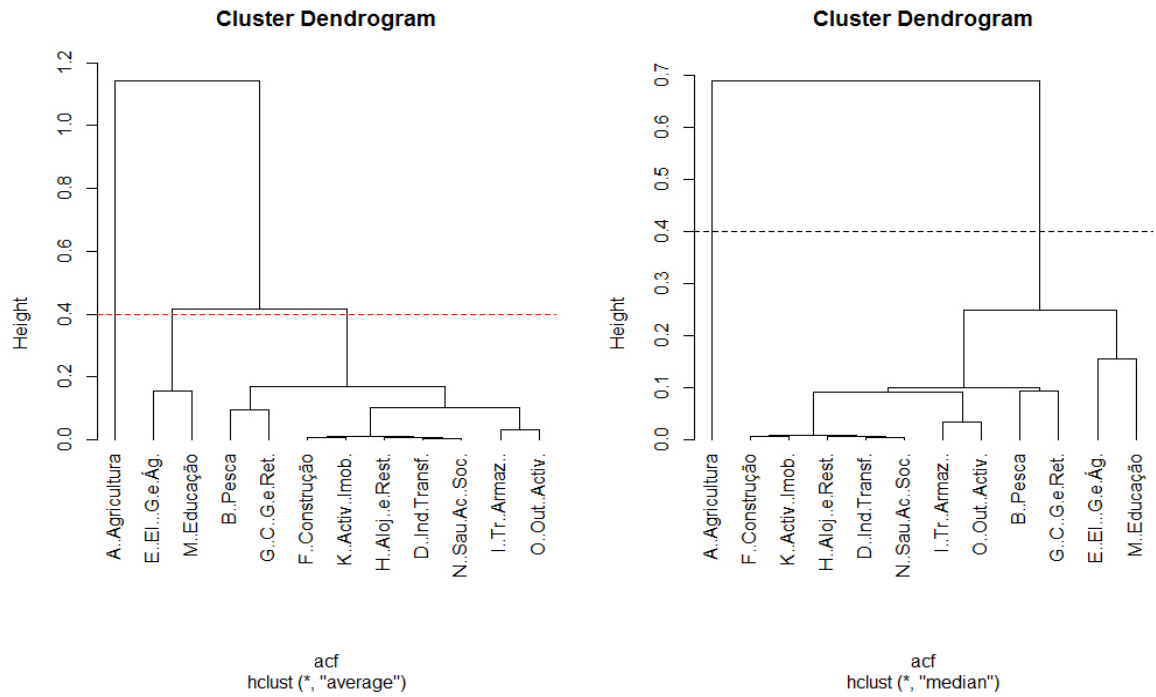


Figura 11. Dissimilaridade baseada no periodograma integrado (método da média e método McQuitty)

